

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 25, 2004, 01:21:23 ; Search time 1098 Seconds

(without alignments)
8264.254 Million cell updates/sec

Title: US-10-010-720-14

Perfect score: 10812

Sequence: 1 MSGGAEMKXSTPGSLFLSP.....NISNLKXISNPPGSLRTT 2136

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delpext 7.0	

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPRO.spool/US10010720/runat_23092004_154319_25232/app_query.fasta_1.2311
-DB=N-Geneseq_29Jan04 -OPMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOEM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US10010720 @CGN 1.1 818 -runat_23092004_154319_25232 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N-Geneseq_29Jan04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002s:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10810	100.0	6411	8	ACH03770 DNA encod
2	10794.5	99.8	7280	6	AAD34306 Human PKI
3	10753.5	99.5	6690	8	ACH03767 DNA encod
4	10677	98.8	7149	8	ABZ59199 Human WNK
5	10677	98.8	7149	8	ACH03764 DNA encod
6	10646	98.5	6327	8	ACH03778 DNA encod
7	10646	98.5	7152	9	ADG99101 Human KRP
8	10589.5	97.9	6606	8	ACH03775 DNA encod

9	10513	97.2	7065	8	ACH03772 DNA encod
10	10508	97.2	6231	8	ACH03786 DNA encod
11	10451.5	96.7	6510	8	ACH03783 DNA encod
12	10375	96.0	6969	8	ACH03780 DNA encod
13	10344	95.7	6147	8	ACH03794 DNA encod
14	10287.5	95.1	6426	8	ACH03791 DNA encod
15	10211	94.4	6885	8	ACH03788 DNA encod
16	9929	91.8	6000	8	ACH03771 DNA encod
17	9872.5	91.3	6279	8	ACH03768 DNA encod
18	9796	90.6	6738	8	ACH03765 DNA encod
19	9765	90.3	7518	8	ACH03779 DNA encod
20	9729.5	90.0	7328	4	AA44683 Novel pro
21	9708.5	89.8	6195	8	ACH03776 DNA encod
22	9632	89.1	6654	8	ACH03773 DNA encod
23	9627	89.0	5820	8	ACH03787 DNA encod
24	9570.5	88.5	6099	8	ACH03784 DNA encod
25	9494	87.8	6558	8	ACH03781 DNA encod
26	9463	87.5	5736	8	ACH03795 DNA encod
27	9406.5	87.0	6015	8	ACH03792 DNA encod
28	9330	86.3	6474	8	ACH03789 DNA encod
29	4053	37.5	2949	8	ACH03766 DNA encod
30	4043	37.4	2490	8	ACH03769 DNA encod
31	3889	36.0	2865	8	ACH03774 DNA encod
32	3879	35.9	2406	8	ACH03782 DNA encod
33	3751	34.7	2769	8	ACH03785 DNA encod
34	3741	34.6	2310	8	ACH03793 DNA encod
35	3587	33.2	2685	8	ACH03790 DNA encod
36	3577	33.1	2396	4	AA157893 Human pol
37	3552.5	32.9	2663	4	AA157893 Human pol
38	3498	32.4	2663	4	AA157893 Human pol
39	2953	27.3	2310	5	AAH76212 Human kin
40	2859.5	26.4	19167	4	AA15340 Human rep
41	2859.5	26.4	19167	4	AB198209 Human tes
42	2859.5	26.4	19167	5	AB19101 Human ner
43	2642	24.4	7825	7	ACC57741 Human pro
44	2627	24.3	1536	4	AA159679 Human pol
45	2627	24.3	1536	4	AA159678 Human pol

ALIGNMENTS

RESULT 1
ID ACH03770 standard; DNA; 6411 BP.

XX ACH03770;

XX AC 26-SEP-2003 (first entry)

XX DE DNA encoding novel human protein kinase #7.

XX KW Human; kinase; ds; gene; cosmetic application; nutraceutical application.

XX OS Homo sapiens.

XX OS US6541252-B1.

XX PD 01-APR-2003.

XX PF 14-MAY-2001; 2001US-00854856.

XX PR 19-MAY-2000; 2000US-0206015P.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Walke DW, Hilbun E, Donoho G, Turner CA;

XX DR WPI; 2003-575927/54.

XX DR P-PSDB; ABO44387.

XX PT New nucleic acid encoding novel human proteins, useful in cosmetic and nutraceutical applications.

QY 621 AlaSerValSerThrGlnValGluProGlnGluProGlnAlaAspGlnHisGlnGlnLeu 640
DB 1861 GCTTCAGTTCTTCAACAGTAGAACCTGGAAGAACCTGAGCAGATCAACATCAACAACCTA 1920
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
DB 1921 CAGTACCAAGCAACCAAGTATATCTGTGTATCTGATGGAGCGGTGACATGTGTCAAGGGA 1980
QY 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
DB 1981 TCTCTGTCTTCCAGCAATCTCCAGTACAGCAAGCCCAACAGCAAGTTTCAATGTGTTCCCA 2040
QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
DB 2041 CATTGAACAGGCACTTTTACAGGACAGCTCCAGGGCATATACCTTCTACTGTCCAAAGCA 2100
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
DB 2101 CAGTCTCAGCCCAATGGGGTATATCCACCTCAAGTGTGGACAGGGGCGAGGCGAGGGT 2160
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
DB 2161 CAGCATCTTCAAGTAGCTTAAACAGGGGTTTCACTTCCCAACCATTAACATCTCTCAG 2220
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
DB 2221 CAGAGAGAGGGAATACAGCAGACAGCCCTCTCAACAGACAGTATTCACCTTCA 2280
QY 761 GlnThrSerThrSerSerGlyAlaIleThrThrAlaGlnProValSerGlnProGlnAlaPro 780
DB 2281 CAGCATCAACCTTCCAGTGAAGCCACATCTGCACAGCAGTGAATCTCAGCTCAAGCTCCA 2340
QY 781 GlnValLeuProGlnValSerAlaGlyValGlnSerThrGlnGlyValSerGlnValAla 800
DB 2341 CAAATCTTGCCTCAAGTATCAGCTGGAAGAACAGAGTCTCAGGGAGTCTCAGGTTGCT 2400
QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
DB 2401 CCTGCAGAGCCAGTTGACAGTAGCAACAGCCCAAGCTTCCAGCGACACACTTGGCTTCC 2460
QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGlnAsnVal 840
DB 2461 TCTGTACAGAGTGCACATTCAGATGCTTCAAGTATGAGTGTGCAATAGCAACAGCAGTC 2520
QY 841 ProSerSerSerGlyAspGlnGlyValArgThrThrIleAspGlnHisTyrArgGlySerVal 860
DB 2521 CCATCTTCCAGTGAAGGACATGAAGGAAGAACTCAAAACGCGCATTAACGAAATCTGTA 2580
QY 861 ArgSerArgSerArgHisGlnGlyLeuThrSerArgProIleAspGlnLeuAsnValSer 880
DB 2581 AGGAGTCCCTCTCCAGCATGAAGAAAACCTTCAAGCCCAAAATTAAGAAATTTGAAATGTTCA 2640
QY 881 AsnIleGlyAspAspArgValAlaGluCysGlnLeuGluThrHisAsnIleGlyMetValThr 900
DB 2641 AATAAAGAGAGACCAAGTAGTAGAATGTCAATTAAGGACTCATTAATAGAAAATGTTACA 2700
QY 901 PheIlePheAspLeuAspGlyAspAsnProGlnGluIleAlaThrIleMetValAsnAsn 920
DB 2701 TTCAAATTTGACCTTAGAGTGTGACAAACCCAGAGGATGACAAACATTAATGTGTGACAAAT 2760
QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
DB 2761 GACTTTATTTCTAGCAATGAGAGAGAGTCTTTGTGATCAAGTCCGAGAAATTAATTTAA 2820
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGlnGlyAspGlnGlyLeu 960
DB 2821 AAAGCTATGAATATCTCAGTAGAGATGTCAGTGTGAAACCAAGAGGATGATCAGGATG 2880
QY 961 GlnSerLeuGlnGlyValAspAspPyrGlyPheSerGlySerGlnIleLeuGlnGlyGlu 980
DB 2881 GAGAGTCTTACAAAGGAAGATGACTATGGCTTTTCAAGTCTTCAAAATTTGAAAGAGAG 2940
QY 981 PheIleGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000

DB 2941 TTCAAACCAACCAATTCCTCGCTTCCATGACCAACAAATAGCAATTCCTTACAGTCT 3000
QY 1001 LeuThrGlnValAlaHisSerAlaGlyArgAspPheIleValSerProValProGluSer 1020
DB 3001 TTTACTCAAGTTCTTATCTCTGCGGGAAGCGGTTTATATGATGATCTGTGCGCAAGAAAGC 3060
QY 1021 ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
DB 3061 CGATTACGAAATCAAAAGTTTCCCGAGTGAATTAAGATACAGTGTGCTGCTCTTACA 3120
QY 1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAla 1060
DB 3121 GCTCAGAGCCCTGGAATGAATGTGTCACTGTGATCAATCCTTATGTTTACAAAGGCC 3180
QY 1061 PheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsnThrAlaProAsnPhe 1080
DB 3181 TTTTCTGAATTTGAGTGTGCCCAATGACAGAAAGAACCCAAACAGCACTCCAACTTT 3240
QY 1081 SerHisThrGlyProThrPheProValAlaProPheLeuSerSerIleAlaGlyVal 1100
DB 3241 AGTCATACAGGACCAACATTTCCAGTAGTACTCTTTCTTAGTAGCATTTGCTGGAGTC 3300
QY 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
DB 3301 CCAACCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3360
QY 1121 SerThrSerValIleGlnSerGluValThrValProThrGlnGlnGlyIleAlaGlyVal 1140
DB 3361 TCCATCATCATTAATTCAGTCTGAGTTACAGTGTCCCAACCAACCTGTGTGAATCA 3420
QY 1141 AlaThrSerThrGlyValAlaThrSerGlyGlyLeuProIleProProValSerGluSer 1160
DB 3421 GCCACCAAGCAGAGTGTGTAACTTCAAGTGTGTCCCAACCAACCTGTGTGAATCA 3480
QY 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValAlaSerIleSerThr 1180
DB 3481 CCAATCTTCCAGCGTAGTTTCAAGTATCAATACATACATCTGCAAGTGTCAATATCACT 3540
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGlnIleValAlaSerSerThrAla 1200
DB 3541 ACATCCCGTCACTTCAAGTGTCCCAATCCCAATCCCAATGATCTGTGTTTCTAGTAGCA 3600
QY 1201 LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
DB 3601 CTGTATCTTCAAGTACAGTTTGCACAACTTCAAGTGTGTGAGGAGGAGTGTGATACC 3660
QY 1221 ProGlyProLysProProAlaValAlaSerGlnGlnAlaAlaGlySerThrValGly 1240
DB 3661 CCAGTCTCTAAGCCTCCAGCTGTAGTATCTCAGAGGACAGCAGGACACTACTGTGGA 3720
QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
DB 3721 GCCCATTAACATAGATTTCTACACCACTTCACTTCCAGCAGCAGCAGTCCACAGTGTCC 3780
QY 1261 IleGlnLeuSerSerSerThrSerThrProThrLeuAlaGluThrValValSerAla 1280
DB 3781 ATTACAGTTAGCAGACAGTACTTACTCTCTTACTTACTTAAAGCGTGTAGTACGCA 3840
QY 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
DB 3841 CACTCAGTGAATGAACATCTCATAGCAGTCAACTGATGATGCTTCTCTCTCTCTCA 3900
QY 1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerTyrIleSerGlnProGlyIle 1320
DB 3901 CCATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3960
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
DB 3961 CTGATCTCTTGTGATTCATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 4020
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360

Db 4021 GGACCTACTTCTACACCTTTATTATCCCAAGTACCTAGTATCCACCTTGATACAGCCT 4080
QY 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 4081 GTTGCACATGTCCTGCTGATACAGACAACTAATTCATATGACACCTCCACACCTTTG 4140
QY 1381 LeuProAsnGlnProHisThrHisCysProGlnValAspSerAspThrGlnProHisAla 1400
Db 4141 CTTCACCAACAGCCCACTACTCATTTGTCTCTGAAGTATGATTCGATACACCAACCAAGCT 4200
QY 1401 ProGlyIleAspAspIleLysThrLeuGlnGlnLysLeuArgSerLeuPheSerGlnHis 1420
Db 4201 CCGTAATGATGACATAAAGCTCTAGAGAAAGAGCTGCGGTCTCTGTTCAGTGAACAC 4260
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
Db 4261 AGCTATCTGGAGCTCAGCATGCTCTCTGTCTCAGAGACCTCACTAGTCAATAGACAGC 4320
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db 4321 ACTGTCAACACAGGATCCCACTACTGCTGTGACCAAGCAAACTCTGACTTCTAC 4380
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4381 ACAAGTACTTCTTACCAACCAACCAATTTACCACTAGGACAGATGCTTGTGCCAGTTACA 4440
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
Db 4441 CCAGTGTGCACACCTGGGCAAGTTCTACCCAGTCAACCTACTACATCAAGAGATGAAA 4500
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValIleLeuProValGlyThr 1520
Db 4501 CCGTAAGTCTGCTCCCTCCACAGCCACTCTAAGAGTCCGGGTGCTGCCAGTGGTACT 4560
QY 1521 GlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu 1540
Db 4561 GAACCTTCAGAGGATCTTACCCAGGACAGGCTGCCACCTTTCCAGAGACCTTCTCTA 4620
QY 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgArgThrLeuSerPro 1560
Db 4621 ACCGAGTCCACGCACTCTAGAGATCTTATGCTCAATTTGAGAAAGAACCTTAGTCCA 4680
QY 1561 Glu***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4681 GAGATKATCAAGAGACTTCTGCGGTGCTGCTGTGTCAGGGGCTCCAAAGCATC 4740
QY 1581 ThrGlnValGlyThrGlnProGlnLysGlyValSerGlnValLysGlnGlyProValLeu 1600
Db 4741 ACAGAAACAGAAACAGGCTCAGAGAGGTCTTCTCAAGTCAAAAGAGGCCCTGCTCTA 4800
QY 1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db 4801 GCACATGATTCAGAGAGCTGGGTGTTTAAATGGAAGCATTTCAAGTTCTGTTGACGA 4860
QY 1621 AspGlyValaGlnLysGlnGlyLysAsnLysSerGlnAspAlaLysSerValHisPheGln 1640
Db 4861 GACGGTCCCAAGAAAGGGGTAAATAATAGTCAGAAAGCATGCAAGTCTGTTTGA 4920
QY 1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValLys 1660
Db 4921 TCCAGGACCTCAGAGTCTCAGTCAATCAAGTACAGTCAAGTACTCTTGCTGTA 4980
QY 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db 4981 CCAGAGCCGATAGGATACATCCCTGATATCTTTCAATGTCAGAGAGGAGGCCAC 5040
QY 1681 LysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnProThrLysValGlyArgPhe 1700
Db 5041 AAAAATACTAGCTCAGAGGCAAAAGTGAACACTGGGAGCTTACCAAGTGTGAACGTTT 5100
QY 1701 GlnValThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnAspLys 1720
Db 5101 CAGGTGACAACTACACCAAAAGAGGTGGTCTTCTGTATCAAAAAGTGTGAGCAAG 5160

QY 1721 IleThrAspThrLysLysGlnGlyProValAlaSerProProPheMetAspLeuGlnGln 1740
Db 5161 ATCACTGACACAAAGAAAGAGACCAAGTGCATCTCTCTTATGATTTGAGACAA 5220
QY 1741 AlaValLeuProAlaValIleProLysLysGlnLysProGlnLeuSerGlnProSerHis 1760
Db 5221 GCTGTTCTTCCGCTGATATCCAAAGAAAGAGAGCTGTAAGTGTACAGAGCTTCACAT 5280
QY 1761 LeuAsnGlyProSerSerSerAspProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db 5281 CTAATAGGCGCGTCTTTCGACCCGAGGCCCTTTTAAAGTAGGATGTGATATGCT 5340
QY 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
Db 5341 TCCGATATCCACACTCGCCCATCAGCTGAGTCAAAAGAGCTTCTTACGACAGATCTA 5400
QY 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerTyrtMetSerSerAspAsnGlnSerAsp 1820
Db 5401 AGTCAAGCCCTTAGTATTCATTTAACTCTCTTACATGATGATGAGCAATGATGACAGAT 5460
QY 1821 IleGlnAspGlnAspLeuLysLeuGlnLeuArgArgLeuArgAspLysHisLeuLysGln 1840
Db 5461 ATCGAAGATGAGACTTAAGTTAGACTGCGACGACTACAGATTAACATCTCAAGAG 5520
QY 1841 IleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLeuTyrtThrLysLeuGly 1860
Db 5521 ATTGAGACCTGACAGATGCGCAGAGACATGAAATTAATCTTTGATACCAAACTGGGC 5580
QY 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgArg 1880
Db 5581 AAGTGTCCCTGCTGTTATTTATTTCCCAAGCTGCTCCCTTCAAGAGAGAGAGCA 5640
QY 1881 ProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerPro 1900
Db 5641 CCCACTAAAGCAAAAGCAGAGAAATCTAGTCAAGAGAGTCTCTGGGGAATTAAGCCCC 5700
QY 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
Db 5701 CAGCTTTCAGATACCTGCTGTGTCAGAGTCAAGCTTCAAGTCTTGCACCCCCAGCAAGC 5760
QY 1921 LeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGlnLeuGlnProLeuLys 1940
Db 5761 CTCCACCTCTCTGCAACATCCAGAGTCCGGGCGAGATAGCTGTACAGCCCCCTTAAG 5820
QY 1941 ProSerProSerSerAspAsnLeuTyrtSerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db 5821 CCATCTCCCTCCAGTACCACTCTATTCAGCTTCCAGGTATGATGTCATTTCAAGTA 5880
QY 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db 5881 CCAAGCTTTCCTCTCCAGGTCAAGAGAACAGGACCAAAACCTGTTGGGGCAACAGT 5940
QY 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
Db 5941 AACAGCCAGCCGCCCAAGTCAAGCTCTGCTCATGATGCTCCAGACAGAGAGGACATTC 6000
QY 2001 ThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGly 2020
Db 6001 ACAGATGACTTGCACAAAGTTGTAGCAATTTGGGCCGGAATGCCATGATCTCTAGGC 6060
QY 2021 ArgArgGlySerLysGlyHisMetAsnTyrtGlnGlyProGlyMetAlaArgLysPheSer 2040
Db 6061 AGGAGAGGAAAGCAAAAGGACATGAATTTAGAGGGCCCTGGAATGCAAGAGATCTCT 6120
QY 2041 AlaProGlyGlnLeuGlyLysIleSerMetThrAsnLeuGlnGlyLysSerAlaProIleSer 2060
Db 6121 GCACCTGGGCAATGTGATCTTCATGACCTGCAACCTGGGTGCTCTGCCCATCTCT 6180
QY 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnTyrt 2080
Db 6181 GCAGCATCAGTACCTCTCTAGGTCACTTCAACCAATGATATGCCCCCAAGCAAGCATAT 6240

QY 2081 GlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThrGlyGlyProAlaProGln 2100
 DB 6241 GGCTTTCAGCTACCCCATTTGGGCTCAATGAGTGAGGAGCGGTGGCCAGCACACAG 6300
 QY 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAenPheAenLleSerAen 2120
 DB 6301 CCACCTGGCCAGCTTCCAACTGGGAACTGCTCCCTTCGAGATTTCACATCAGCAAT 6360
 QY 2121 LeuGlnLysSerLleSerAenProGlySerAenLeuArgThrThr 2136
 DB 6361 TTGCAAAATTCATCAGCAACCCCAAGCTCCAACTCGAGCACT 6408
 RESULT 2
 AAD34306
 ID AAD34306 standard; cDNA; 7280 BP.
 XX
 AC AAD34306;
 XX
 DT 16-JUL-2002 (first entry)
 XX
 DE Human PKIN-9 cDNA.
 XX
 KM Human; kinase; enzyme; PKIN-9 protein; immune system disorder; anaemia;
 KM acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;
 KM aschma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
 KM AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
 KM leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
 KM Down's syndrome; gene therapy; protein therapy; cytostatic; gene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 419..6826
 FT /tag= a
 FT /product= "Human PKIN-9 protein"
 FT
 XX WO200218557-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 31-AUG-2001; 2001WO-US027219.
 XX
 PR 31-AUG-2000; 2000US-0229873P.
 PR 08-SEP-2000; 2000US-0231357P.
 PR 14-SEP-2000; 2000US-0232654P.
 PR 22-SEP-2000; 2000US-0234902P.
 PR 29-SEP-2000; 2000US-0236492P.
 PR 06-OCT-2000; 2000US-0238389P.
 PR 13-OCT-2000; 2000US-0240542P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Bandman O, Nguyen DB, Walla NK, Hafalia AJA, Yao MG, Gandhi AR,
 PI Gururajan R, Ding L, Patterson C, Yue H, Baughn WK, Tribouley CM,
 PI Thornton M, Elliott JD, Ison CH, Au-Yang J, Tang YF,
 PI Azimzai Y, Burdill VS, Marcus GA, Zingler KA, Lu DM, Lal PG,
 PI Rankumar J, Warren BA, Kearney L, Policky UL, Thangavelu K,
 PI Burford N;
 XX
 DR WPI; 2002-329769/36.
 DR P-PSDB; AAE21714.
 XX
 PT New human kinases, useful for diagnosing, treating or preventing immune
 PT system disorders (e.g. Crohn's disease) neurological disorders (e.g.
 PT epilepsy), or cell proliferative disorders (e.g. cancers such as leukemia
 PT or lymphoma).
 XX
 PS Claim 88; Page 203-205; 218pp; English.
 CC The present invention relates to human kinases (PKIN) and polynucleotides
 CC encoding such proteins. PKIN sequences of the invention are useful for
 CC diagnosing, treating or preventing disorders associated with aberrant
 CC expression of PKIN, particularly immune system disorders (e.g. acquired

CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
 CC anaemia, aschma, neurological disorders (e.g. epilepsy, Charcot-Marie-
 CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
 CC such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma),
 CC and developmental disorders (e.g. Down's syndrome). They are also used in
 CC gene therapy and protein therapy. The present sequence is a cDNA encoding
 CC human PKIN-9 protein
 XX
 SQ Sequence 7280 BP; 1940 A; 1999 C; 1720 G; 1621 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 7280
 Score: 10794.50 Matches: 2134
 Percent Similarity: 99.91% Conservative: 0
 Best Local Similarity: 99.91% Mismatches: 1
 Query Match: 99.84% Indels: 1
 DB: 6 Gaps: 1
 US-10-010-720-14 (1-2136) x AAD34306 (1-7280)
 QY 1 MetSerGlyGlyAlaAlaGluLysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 DB 419 ATGCTTGGCGCGCCGCGAGAGACAGACAGCACTCCGGTTCCTCTCTCGCG 478
 QY 21 ProAlaProAlaProLysAenGlySerSerSerSerValGlyLysLeuGly 40
 DB 479 CCGGCTCTGCGCCCAAGATGCTCCAGCTCCGATTCCTCGGGGAGAAACTGGGA 538
 QY 41 AlAlAlAlAlAlaAlaAlaValThrGlyArgThrGluGluLysArgArgArgHisThr 60
 DB 539 GCCCGCGCGCCGCGCTGTGACCGGACGACGAGAGTACAGCGCGCGCCGACACT 598
 QY 61 MetAspLysAspSerSerValGlyAlaAlaAlaThrThrThrThrThrGluHisArgPhePhe 80
 DB 599 ATGGAACAAGACAGACCGCTGGGGCGCGCCGACCACTACCACTGAGCACTGCTTTC 658
 QY 81 ArgArgSerValLleCysAspSerSerAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 DB 659 CGCGGAGCGCTCATCTCGACTCCCATGCACTGAGAGCTTCCGCGCTTCTCTT 718
 QY 101 SerLeuProGlnProSerLleProAlaAlaValProGlnSerAlaProProGluProHis 120
 DB 719 TCCCTGCCCCAGCCACATCCCGCGCGTGTCCGAGAGTGTCCACGAGCCGAC 778
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 DB 779 CGGAAAGACCGTGAACCGCCACCTCCAGGTAGCCAGACCTCCAGCGCT 838
 QY 141 AlAlAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 DB 839 GCGCGCCCTGGGAGAACAGCGCGTGGCGGCTCCCTCGACTGTCTCCAGCAGTACC 898
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProProPro 180
 DB 899 AGCAAAAGACCGCCAGTGTCCAGCTTGTGGGAGCAAAAGAGACCCCGCG 958
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGlnGluGluArgSerGlnGln 200
 DB 959 GCGAGACTGGCAACGCGCGCGGCGAGCCAGGACCAAGAGCAAGAGAAAGAGCAGAG 1018
 QY 201 GlnAspAspLleGluGluLeuGluThrLysAlaValGlyMetSerAenAspGlyArgPhe 220
 DB 1019 CAGATGATATCGAAGAGCTGAGAACCAAGCCGTGGGAATGTCTAACGATGGCGCTTT 1078
 QY 221 LeuLysPheAspLleGluLleGlyArgGlySerPheLysThrValLysLysGlyLeuAsp 240
 DB 1079 CTCAAGTTGACATCGAAATCGAGAGGCTCTTTAAGACGCTTACAAAGGTGGAC 1138
 QY 241 ThrGluThrThrValGluValAlaAlaThrPysGluLeuGlnAenPArgLysLeuThrLysSer 260
 DB 1139 ACTGAACCAACCGTGAAGTGGCTGTGTGAATGACGAGATCGAAATTAACAAAGCT 1198
 QY 261 GluArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnLle 280

```
Db 1199 GAGAGCAGAGATTAAAGAGAGCTGAAATGTTAAAGCTTCAGATCCAAATATT 1258
Qy 281 ValArgPheTyrAspSerTyrGluSerThrValLysGlyLysCysIleValLeuVal 300
Db 1259 GTTAATTTTATGATTCCTGGGAATCCACAGTAAAGAGAAAGAGATGCTTTTGGG 1318
Qy 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Db 1319 ACTGAACTTATGACCTCTGGAAACCTTAAACGTTATCGAAAAGGTTTAAAGTGAATGAG 1378
Qy 321 IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
Db 1379 ATCAAGATTCTAAGAGCTGGGCGCTGAGATCTTAAAGGCTTCAGTTCTTCATCT 1438
Qy 341 ArgThrProProlIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 1439 CGAACTCCACTATCATTCACCGCGATCTTAATGTGACAAATCTTTATCACCGGCTT 1498
Qy 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
Db 1499 ACTGGCTCAGTCACAGATTGGAGACCTCGGTCTGGCAACCTGAAGCGGCTTCTTCC 1558
Qy 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
Db 1559 AAGAGTGTAGTAGTATCCCGAGAGTTCAAGCCCTGAGATGTATGAGAGAAATATGAT 1618
Qy 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db 1619 GAATCCGTGAGCTTATGATCTTTGGAGTGTGACATGCTTGAGATGTGATCTGAATAT 1678
Qy 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
Db 1679 CCTTACTCGAGGTGCGAAAATGCTGCGCAGATCTACCTCCCGTCAACGAGCGGCTGAG 1738
Qy 441 ProAlaSerPheAspLysValAlaIleProGluValLysGlnIleIleGluGlyCysIle 460
Db 1739 CCAGCCAGTTTTCAGAAAGTGCAAATTCCTGAAGAGAAATATATGAGAGATGCATA 1798
Qy 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuAsnHisAlaPhePheGln 480
Db 1799 CGAAMAAACAAAGATGAAGATATTCATCAAAAGCCTTGAACCAATGCTTCTTCAA 1858
Qy 481 GlnGluThrArgValArgValGluLeuAlaGluLysAspGlyGluLysIleAlaIle 500
Db 1859 GAGAAAACAGAGTACGGGTAGATTTAGCAGAAAGAGTGTGAGAAATTAAGCCATA 1918
Qy 501 LysLeuTyrPheLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
Db 1919 AAATTTATGGCTACGATTTGAAGATTTAAGAAATTTAAAGGAAATACAAAGATATGAA 1978
Qy 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 1979 GCTATTAGTTTCTTTGATTATTAAGAGAGATGCCCCAGAAAGATGTTGCACAAAGAAATG 2038
Qy 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 2039 GTAGAGCTGGGTATGCTGTGAAGGTGATCAAAAGCCATGGCTTAAAGCATCAAAAGAC 2098
Qy 561 ArgValSerLeuLysArgLysArgGluGlnArgGlnLeuValArgGluGlnGlu 580
Db 2099 AGAATATCATTTAATTAAGAGAAACAGAGACGCGGATGTGTACGGAGAGCAAGAA 2158
Qy 581 LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGlnGlnSerSerAlaSerGln 600
Db 2159 AAAAAAAGCAGAGAGAGACAGTCTCAAAACAGAGGAGAGAACATCCAGTCTTCCAG 2218
Qy 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 2219 ACAGAGATCAAGACGCTCCCTTCTGTAGACACCGGCACTACTACTTCTACCACTTCA 2278
Qy 621 AlaSerValSerThrGlnValGlnProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db 2279 GCTTCAGTTTCTACACAAAGTGAACCTGAAAGAACTTGAAGCATCAATCAACACTA 2338
Qy 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 2339 CAGTACAGACCAACCGATATATCTGTATTATCTGATGAGCGGTTGACAGTGTGAGGGA 2398
Qy 661 SerSerValPheThrGluSerArgValSerSerGlnGlnIleValSerTyrGlySerGln 680
Db 2399 TCTCTGCTTTCACAAATCTCGAGTGTGAGCCAAACAGACAGATTCATATAGTTCCTCAA 2458
Qy 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2459 CATGAACAGGACATTTCTACAGGACAGATCCAGGAGCATATACCTTACTCTCCAAACA 2518
Qy 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 2519 CAGTCTCAGCCCAATGGGATATATCCACCCCAAGTGTGGACAGGGGCGAGCGCAGG 2578
Qy 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2579 CAGCCATCTCAAGTACTTAAACAGGGGTTTCACTTCCAAACCAATCAACATCTCT--- 2635
Qy 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2636 CAGCAGCAGGGAATACAGACAGACAGCCCTCTCAACAGACAGTGCAGATTCACCTTCA 2695
Qy 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2696 CAGACATCAACCTCAGAGAGCCCACTGACACAGCCAGTGAATCAACCTCAAGCTCCA 2755
Qy 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlyValAla 800
Db 2756 CAAGTCTTGGCTCAAGTATCACTGAGAAACAGAGTACTCAGGAGTCTTCAAGGTGCT 2815
Qy 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db 2816 CCTCAGACAGCCAGTGCAGTGCAGACAGCCCAAGCTCCAGCCGACACACTTGGCTTCC 2875
Qy 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
Db 2876 TCTGTAGACAGTGCACATTCAGATGTGCTTCAAGTATGAGTATGAGCAATGAAAGAGTC 2935
Qy 841 ProSerSerSerGlyArgHisGluGlyArgThrThrLysArgHisTyrArgLysSerVal 860
Db 2936 CCATCTTCCAGTGAAGAGCATGAAGGAAAGAACTCAAAACGCGCTTACCGAAATCTGTA 2995
Qy 861 ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
Db 2996 AAGAGTGGCTCTCGACATGAATAAACTTCAAGCCCAAAATTAAGAAATTTGAATGTTCA 3055
Qy 881 AsnLysGlyAspArgValAlaGluCysGlnLeuGlnLysHisAsnArgLysMetValThr 900
Db 3056 AATTAAGAGAGCCGAGTAGTAGAATGTCAATTTAAGACTCAATATATGAGAAATGTTACA 3115
Qy 901 PheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAlaThrIleMetValAspAsn 920
Db 3116 TTCAAAATTTGACCTAGATGTGACAAACCCGAGAGATTCAGCAACAATTTAGTGAACAT 3175
Qy 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
Db 3176 GACTTTATTTCTAGCAATAGAGAGAGAGTCTGTTGTGATCAAGTGCAGAAATTAATGAA 3235
Qy 941 LysAlaAspGlnMetLeuSerGluAspValSerValGlnProGluGluLysAspGlnGlyLeu 960
Db 3236 AAAGCTGATGAATATGCTCAAGTAGAGATGCTAGTGTGAACCAAGAGGATGATCAGGGATG 3295
Qy 961 GluSerLeuGlnGlnLysAspAspTyrGlyPheSerGlySerGlnLysLeuGluGlyGln 980
Db 3296 GAGAGTCTCAAGAGAAAGATGACTATGCGCTTTTCAAGTCTCAAGAAATTAAGAGAGAG 3355
Qy 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
Db 3356 TTCAAACAAACAAATTCCTGCGTCTTCCATGCAACAGCAAGAAATAGCAATTCCTACAGTCT 3415
```

QY 1001 LeuThiGlnValAlaHisSerAlaGlyArgPheIleValSerProValProGluSer 1020
Db 3416 TTAACTCAAGTTGTTCTTCGGCGGAAGGGGTTTATGAGTCTCTGGCCAGAAAGC 3475
QY 1021 ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaIleSerThr 1040
Db 3476 CGATTACGAAATCAAAAGTTTCCCGAGTAATAACAGTACAGTGGCTGCTTACA 3535
QY 1041 AlaGlnSerProGlyMetLeuLeuSerHisSerAlaSerSerLeuSerLeuGlnAla 1060
Db 3536 GCTCAGAGCCCTGGAAAGAACTTCTCAGCTCTGCAATCATCTTAGCTTACCAACAGGCC 3595
QY 1061 PheSerGluLeuArgAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe 1080
Db 3596 TTTTCTGAACTTAACGTGGCCAAATGACGAAGAACCCAAATACAGACCTCCAAATTT 3655
QY 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal 1100
Db 3656 AGCATACAGGACCAACATTTCCAGTAGTACCTCTTCTTAAGTACATTTGCTGGAGTC 3715
QY 1101 ProThrThrAlaAlaIleThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
Db 3716 CCAACACAGCAGCAGCAGCAGCAGCTCTGCAACAGAGCCCTCTCAATGACATTT 3775
QY 1121 SerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGlyVal 1140
Db 3776 TCCACATCAGTAATTCAGTCTGAGGTTACAGTCCCACTGAAAGAGGGAATTCGCTGAGTT 3835
QY 1141 AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer 1160
Db 3836 GCCACCCAGCAGAGTGCGTAATCTCAGTGCGTCCCAACCACTGGTGCTCAATCA 3895
QY 1161 ProValLeuSerSerValIleSerSerIleThrIleProAlaValIleSerIleSerThr 1180
Db 3896 CCACTACTTCCAGCGAATTCAGATTCACAAATACCTGACAGTGTGCTCAATATCACT 3955
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValIleSerSerThrAla 1200
Db 3956 ACATCCCGCTCACTTCAAGTCCCAACATCCACATCGATGATGTTTCTAGTACAGCA 4015
QY 1201 LeuThrProSerValIleThrValSerAlaThrSerAlaSerAlaGlyIleSerThrAlaThr 1220
Db 4016 CTGTATCTTTCAGTACAGTTCAGCAATTCAGCTCTGCAAGGCGCAGACTGCTAAC 4075
QY 1221 ProGlyProLysProProAlaValIleSerGlnGlnAlaGlySerThrThrValGly 1240
Db 4076 CCAAGTCTTAAGCTCCAGCTGATGATCTCAGACGACGACGACACTACTGTGAG 4135
QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 4136 GCCACATTTAATCATGATTTCTACACACACTTCATTCACAGACAGCTCCACACTGCTC 4195
QY 1261 IleGlnLeuSerSerSerThrThrThrProThrLeuAlaGluThrValValIleSerAla 1280
Db 4196 ATTACAGCTTACACAGACTTCTACTCTTCTTAGCTGAACCGTGGTGTAGTACGCCA 4255
QY 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 4256 CACTCAGATGAAGACATCTCATAGACAGTACAACTGGATTGGCTTCTCCCTCTCGCA 4315
QY 1301 ProSerSerSerSerSerProGlyValAlaGlyValSerSerThrIleSerGlnProGlyGly 1320
Db 4316 CCACTTCTCTCTCTCTCTCGGAGCAGAGAGTGTCTAGTTATTTCTCCACCTCGTGGG 4375
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
Db 4376 CTGATCTCTTGGTATTCATCAGTATGATCTTCTACTCTATTTCTTCCCAAGCAGCA 4435
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGluPro 1360
Db 4436 GGACCTACTTCTTACCTTTATTAACCCAGTACCTAGTATCCACCTTGTTGACAGCT 4495
QY 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 4496 GTTGCCAAATGCTGCTGTGTACAGACAGACTAATTCATGATGAGCTTCAACAGCTTGG 4555
QY 1381 LeuProAsnGlnProHisThrHisGlySerProGluValAspSerAspThrGlnProLysAla 1400
Db 4556 CTTCCCAACCAAGCCCACTACTCTTCTCGAAGTATGATCTATTAACAAACCCAAAGCT 4615
QY 1401 ProGlyIleAspAspIleLysThrLeuGlnGluLysLeuArgSerLeuPheSerGlnHis 1420
Db 4616 CTTGGAATTTATACATTAAGACTTCTGAAAGAAAAGCTGGGCTCTGTTCAAGTAAAC 4675
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGluSer 1440
Db 4676 AGCTCATCTGAGCTCAGCATGCTCTGTCTCACTGAGACCTCACTAGTCATGAGAGC 4735
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db 4736 ACTGTCAACACAGCATCCCACTACTGCTGTGACCAAGCAAACTCTGACTTCTAAC 4795
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4796 ACAAGTACTTCTTACCAACCAATTTACATAGGAACAGTGTCTTGGCCAGTTACA 4855
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
Db 4856 CCAAGTGTCAACCTGGGCAAGTTTCTACCCAGTCAAGCTACTACATCAGAGAGTAAA 4915
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
Db 4916 CTTGGAATCTGCTCTTCAAGCAACCTCTAATCAAGGCTCGGCTGCGCAAGTGGTACT 4975
QY 1521 GluLeuProAlaGlyThrLeuProSerGluGlnLeuProProPheProGlyProSerLeu 1540
Db 4976 GAATCTCAAGAGTACTTCAACCAAGGACAGCTGACACTTTTCCAGACCTTCTCTA 5035
QY 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgThrLeuSerPro 1560
Db 5036 ACCAGTCCAGCAACTCTTAGAGATCTGATGCTCAATTGAGAAACACTTAGTCCA 5095
QY 1561 Glu***IleThrValThrSerAlaValGlyProValSerMetAlaIleProThrAlaIle 1580
Db 5096 GAGATGATCAAGAGACTTCTGGGTGCTCTGTGTCTCAGTGGGCTCCAAACAGCATC 5155
QY 1581 ThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGluGlyProValLeu 1600
Db 5156 ACAGAGCAGGAACACAGCTCTCAGAGGAGTGTCTCAAGTCAAAAGAGCCCTGTCTTA 5215
QY 1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db 5216 GCAACTAGTTCAGGAGCTGGTGTTTTAATGGAGATTCAGTTTCAGTTTCTGTGACAG 5275
QY 1621 AspGlyAlaGlnLysGluGlyLysAsnLysSerGluAspAlaLysSerValHisPheGlu 1640
Db 5276 GACGCTGCCAGAAAGGGGTAAATAATATATGAAATCAAGATCAAACTGTGTCATTTGAA 5335
QY 1641 SerSerThrSerGluSerSerValLeuSerSerSerSerProGluSerThrLeuValLys 1660
Db 5336 TCCAGACCTCAGAGTCTCTAGTCTTATCAAGTAGTAGTCCAGAGATCACTTGCTGAAA 5395
QY 1661 ProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGluSerAlaHis 1680
Db 5396 CCAAGAGCCGAATGCAATACCATCTCCGATATCTTCAAGATGCGCAAGAGAGGCCAC 5455
QY 1681 LysThrThrAlaSerGluAlaLysSerAspThrGlyGlnProThrLysValGlyArgPhe 1700
Db 5456 AAAACTATGCTTACAGGCAAGTCAACACTGAGCACTGGAGCTTACCAAGGTTGGACGTTT 5515
QY 1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGluAspLys 1720
Db 5516 CAGGTGACAACTACAGCAAAAGTGGTGTCTTCTGTATCAAAAACTGAGAGACAG 5575
QY 1721 IleThrAspThrLysLysGluGlyProValAlaSerProProPheMetAspLeuGluGln 1740


```
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 2341 CAAGCTTGCCCTCAAGATACACTGGAAAAACAGGGCTTCCCACTCGACTGCCACACAG 2400
QY 791 ----- 791
Db 2401 TAACCAAGAGATTCAATATATGTCCTCTTCCAAAGTGCTTGTGTTGATTCATTCCT 2460
QY 791 ----- 791
Db 2461 ACAGTCTTACCCCCCTCCCACTGCCAGACAGAGTACCTGGGTAATCTTCCACA 2520
QY 791 ----- 791
Db 2521 GTGATGACGCTTATGTGAATCAATCTTTAGTTCCTATGGGTGGTGTAGAGAGACAG 2580
QY 791 ----- 791
Db 2581 GTTCAAGTGTCCCAAGCAGAGAGGAGTTTACGACAAAGCCCACTACATCTCCACACA 2640
QY 792 -----SetThrGlnGlyValSerGlnValAlaProAlaGluProValAlaVal 807
Db 2641 GCAGTTTGGAGAGTACTCAGGAGTCTCTCAGGTTGCTCTGACAGCCAGTTCAGTA 2700
QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2701 GCACAGCCCCAAGACTACCCAGCCAGCACTTTGGCTTCTCTGTAGACAGTGCATTTCA 2760
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyYargHis 847
Db 2761 GATGTGCTTACAGATATAGATGATGGCAATGGAAAGTCCCATCTTCCAGTGAAGGCAT 2820
QY 848 GluGlyArgThrThrLysArgHisTyrArgLysSerValArgSerArgHisGlu 867
Db 2821 GAAGGAATACTACAAACCGGCTTACCGAAATCTGTAGAGAGCGCTCTGCACATGAA 2880
QY 868 LysThrSerArgProLysLeuArgLysLeuAsnValSerAsnLysGlyAspArgValVal 887
Db 2881 AAACTTCACGCCCAATTAAGAAATTTGAATGTTCAATTAAGAGAGACCGAGTACTA 2940
QY 888 GluGlyGlnLeuGlnThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly 907
Db 2941 GAATGTCAATTAGAGACTCATTAATAGAAATGTGTACATTCATTAATGACCTTGATGT 3000
QY 908 AspAsnProGlnGluLeuAlaThrLysMetValAsnAspPheLeuAlaLysGlu 927
Db 3001 GACAAACCCCGAGAGATAGCAACAATTAATGTGAACAATGACTTATCTTAGCAATAGAG 3060
QY 928 ArgGlnSerPheValAspGlnValArgGlnLysLeuLysAlaAspGluMetLeuSer 947
Db 3061 AGAGAGTGTGTTGGATCAAGTGAGAAATTAATTAAGAAAAGCTGATGAATGACTCAGT 3120
QY 948 GluAspValSerValGlnProGlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
Db 3121 GAGGATGTCAATGTGGAAACCAAGAGGTGATCAGGATTTGAGAGTCTTCAACAGAAAGAT 3180
QY 968 AspTyrGlnPheSerGlySerGlnLysLeuGlnGlyGluPheLysGlnProLysProAla 987
Db 3181 GACATATGCTTTTCAAGTCTCAAAATTTGAAAGAGAGTTCAACAACCAATTCCTGCG 3240
QY 988 SerSerMetProGlnGlnLysGlyLysProThrSerSerLeuThrGlnValAlaHisSer 1007
Db 3241 TCTTCATGCGACAGCAAAATGAGCATTCCTTCAAGTCTTTTAATCTCAAGTTGTTCAATCT 3300
QY 1008 AlaGlyArgArgPheLeuValSerProValProGlnSerArgLeuArgGluSerLysVal 1027
Db 3301 GCGGGAAGCGGTTTAATAGAGTCTGTGCGAAGAGCCCATTAACAGATCAAAAGTT 3360
QY 1028 PheProSerGlnLeuThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
Db 3361 TTCCCCAGTAAATACAGATACAGTGTGCTCTTACAGCTCAGAGCCCTGGAGATGAAC 3420

QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAla 1067
Db 3421 TTGTCTCACTGTGACATCAATCCCTTAATGTACAAAGGCGCTTTTTCGAATTAAGAGTGC 3480
QY 1068 GlnMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
Db 3481 CAATAGACAGAAAGAGCCCAAAVACAGCACTCCAACTTTAGTCATACAGACCAATTT 3540
QY 1088 ProValValProProPheLeuSerSerLeuAlaGlyValProThrThrAlaAlaThr 1107
Db 3541 CCAGTAGTACCTCTCTTCTTAATAGTACATGTGAGTCCCAACACAGACAGCAGCACA 3600
QY 1108 AlaProValProAlaThrSerSerProProAsnSerPheSerThrSerValLysGlnSer 1127
Db 3601 GCACAGTCCCTGCACACAGAGAGCCCTCTTAATGACATTTCCACATCAGTAAATGACT 3660
QY 1128 GluValThrValProThrGlnGlnGlyLysAlaGlyValAlaThrSerThrGlyValVal 1147
Db 3661 GAGGTTACAGTCCCACTGAAAGAGGAGTTGCTGGAGTTGCCACAGCAGAGTGTGTA 3720
QY 1148 ThrSerGlyGlyLeuProLysProLysProValSerGlnSerProValLeuSerValVal 1167
Db 3721 ACTTCAGTGTGTCCCACTACACACTGTGTGATCAATCAGATCACTTCCAGGCTAGTT 3780
QY 1168 SerSerLysThrLysProAlaValValSerLysSerThrThrThrSerProSerLeuGlnVal 1187
Db 3781 TCAGATACACAAATACCTGCGAGTGTCTCAATATCTTACATCCCGCTCACTTAAGTCT 3840
QY 1188 ProThrSerThrSerGlnLysValValSerSerThrAlaLeuTyrProSerValThrVal 1207
Db 3841 CCCACATCACTGAGATGTTGTTTCTAGTACAGCAGTATCTCTTCACTGAACATCT 3900
QY 1208 SerAlaThrSerAlaSerAlaGlyLysSerThrAlaThrProGlyProLysProProAla 1227
Db 3901 TCACCAACTTACGCTCTGACAGGAGGAGTACTCTACCTCAACCGAGTCTCAAGCTCCAGCT 3960
QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
Db 3961 GTAGTATCTGACAGCAGCAGCAGCAGCACTACTGTGGAGGCCAATTAACATCAGTTTCT 4020
QY 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerLysGlnLeuSerSerThrThr 1267
Db 4021 ACCACCACTTCATTCACAGCAGCAGCTTCAACAGCTGCTCATTCAGCTTAGCAGCAGTACT 4080
QY 1268 SerThrProThrLeuAlaGlnThrValValValSerAlaHisSerLeuAspLysThrSer 1287
Db 4081 TCTACTCTTACTTAGCTGAACCGGTGTAGTACGCCACACTCAAGATTAAGACATCT 4140
QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
Db 4141 CATAGCAGTACAACTGATTTGGCTTTCTCTCTCTGACACATCTTCTCTCTCTCTCT 4200
QY 1308 GlyAlaGlyValSerSerTyrLysSerGlnProGlnGlyLeuHisProLeuValLysPro 1327
Db 4201 GAGACAGAGAGTGTATGTTATTTCTCAGCTGTGGGTGAGCATCTTGTGCTATTTCA 4260
QY 1328 SerValLysLeuAspThrProLysLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
Db 4261 TCAGTATAGCTTACTCTCTATTTCTTCCCAACAGCAGAGACCTTCTTCAACCTTTA 4320
QY 1348 LeuProGlnValProSerLysProProLeuValGlnProValAlaAsnValProAlaVal 1367
Db 4321 TTACCCCAAGTACTATATATCCACCTTGTGTACAGCTGTGTGCAATGTGCTGCTGTA 4380
QY 1368 GlnGlnThrLeuLeuHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
Db 4381 CAGCAGACACTAAATTCATATGACAGCTCAACAGCTTGTCTTCCCAACAGCCCATACT 4440
QY 1388 HisCysProGlnValAspSerAspThrGlnProLysAlaProGlyLysAspAspLys 1407
Db 4441 CATTTGCTTAAGATGATTTCTGATACAAACCAAACTCTGGAATTAAGATGAAGA 4500
QY 1408 ThrLeuGlnGlnLysLeuArgSerLeuPheSerGlnHisSerSerSerGlyAlaGlnHis 1427
```

Db 4501 ACTAGAGAAAAGCTCGGCTCTCTTCAGTGAACACAGCTCATCTGAGCTACGAT 4560
Qy 1428 AlaSerValSerLeuGluThrSerLeuValIleGluSerThrValThrProGlyIlePro 1447
Db 4561 GCGCTGTCTACTGAGACCTCACTAGTCATAGAGACCTGTACACCGAGGATCCCA 4620
Qy 1448 ThrThrAlaValAlaProSerLysLeuLeuThrSerThrThrSerThrCysLeuProPro 1467
Db 4621 ACTACTGCTGTGACCAAGCAAACTCTGACTTCTACCAAGACTTGTCTTACCAACA 4680
Qy 1468 ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValIleThrProGlyGln 1487
Db 4681 ACCAATTACACATAGGAACAGTTGCTTGCAGATTACACAGTGGTCAACCTGGGCAA 4740
Qy 1488 ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys 1507
Db 4741 GTTCTACCCGACGACGACTACTACATCAGAGAGAAACCTGGAACTGTCTCCCTCAG 4800
Qy 1508 ProProLeuThrLysAlaProValLeuProValGlyThrGluLeuProAlaGlyThrLeu 1527
Db 4801 CCACCTCACTAAGAGCTCCGGTGTGCTGCAAGTGGTACTGAACCTTCCAGCAGTACTCTA 4860
Qy 1528 ProSerGluGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnGlnProLeu 1547
Db 4861 CCCAGCGACAGCTGCGACCTTTTCCAGACCTTCTTACCCAGTCCACAGCAACCTCTA 4920
Qy 1548 GluAspLeuAspAlaGlnLeuArgArgThrLeuSerProGlu**IleThrValThrSer 1567
Db 4921 GAGGATCTTGATGCTCAATTGAGAAAGAACCTTATCTCAGAGATATACAGTGCCTCT 4980
Qy 1568 AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGluAlaGlyThrGlnPro 1587
Db 4981 GCGGTGGTCTCTGTCTCCATGGCGGCTCCAAACGCAATCCAGAGCAGAAACAGCT 5040
Qy 1588 GlnLysGlyValSerGlnValLysGlyLysProValLeuAlaThrSerSerGlyValAlaGly 1607
Db 5041 CAGAAAGGCTGTTTCTCAAGTCAAAAGAGCCCTGCTTAGCAACTAGTTCCAGAGCTGCT 5100
Qy 1608 ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyValGlnLysGlnLys 1627
Db 5101 GTTTTAAAGATGGAGCATTTTCAGATTTCTGTGTGAGAGAGCGGCGCCAGAAAGAGGT 5160
Qy 1628 LysAsnLysSerGluAspAlaLysSerValHisPheGluSerThrSerGluSerSer 1647
Db 5161 AAAAATTAAGTCAAGAAAGATGCAAGCTGTCTCATTTTGAATCCAGACCTCAGAGCTCTCA 5220
Qy 1648 ValLeuSerSerSerSerProGluSerThrLeuValLysProGluProAsnGlyIleThr 1667
Db 5221 GTGCTATCAAGTAGTAGTCCAGAGAGTACCTTGTGAAACCAGACCCGAAATGGCTAAAC 5280
Qy 1668 IleProGlyLysSerSerAspValProGluSerAlaHisLysThrThrAlaSerGlnAla 1687
Db 5281 ATCCCTGATATCTTTCAGATGTGCCAGAGAGTGCACAAATCACTAGCTCAAGAGCA 5340
Qy 1688 LysSerAspThrGlnGlnProThrLysValGlyArgPheGlnValThrThrAlaAsn 1707
Db 5341 AAGTAGACACTGGGCGACCTTCAAGAGTGGACCTTTTCAAGTACAACTACAGCAAAAC 5400
Qy 1708 LysValGlyArgPheSerValSerLysThrGluAspLysIleThrAspThrLysLysGlu 1727
Db 5401 AAAGGGGCTGCTTCTCTGATCAAAAACCTGAGCAAGATCACTGACAAAGAAAGAA 5460
Qy 1728 GlyProValAlaSerProProPheMetAspLeuGlnGlnAlaValLeuProAlaValIle 1747
Db 5461 GGACCAAGGCGATCTCTCTCTTAAAGATTGGAAACAAGCTGTCTCTGCTGTGANA 5520
Qy 1748 ProLysLysGluLysProGluLeuSerGluProSerHisLeuAsnGlyProSerSerAsp 1767
Db 5521 CCAAGAAAGAGAAAGCCTGAACTGTCAAGCCTTCAATCTAAATGGGGCGCTCTTGAC 5580
Qy 1768 ProGluAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHisSerPro 1787

Db 5581 CCGAGGCCCGCTTTTAACTAGGAGATGTGCAGTATGTTCCGGTAGTCCAGCTCGCCC 5640
Qy 1788 HisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer 1807
Db 5641 CATCAGCTGAGCTCAAAAGAGCTTCTTACCCAGAAATCTAAGTCAAAAGCTTATGATTA 5700
Qy 1808 PheAsnSerSerLysMetSerSerAspAsnGluSerAspIleGluAspGluAspLeuLys 1827
Db 5701 TTTAACCTCTTTCATAGAGTAGGACATAGTACATGATATCGAAGATGAAGACTTAAG 5760
Qy 1828 LeuGluLeuArgArgLeuArgAspLysHisLeuLysGlnIleGlnAspLeuGlnSerArg 1847
Db 5761 TTAGAGCTGCACACACTACGATTAACATCTCAAGAGATTCCAGACTTCAGAGTGC 5820
Qy 1848 GlnLysHisGlnIleGluSerLeuThrLysLeuGlyLysValProProAlaValIle 1867
Db 5821 CAGAAAGATGAATTTGAATCTTTGTATACCAATCGGCAAGAGTGCCTCTGTATTT 5880
Qy 1868 IleProProAlaAlaProLeuSerGlyArgArgArgProThrLysSerLysGlySer 1887
Db 5881 ATTCGCCAGCTGCTCCCTTTCAGGGAGAGAGACGACGACCACTAAAGCAAGGCGAGC 5940
Qy 1888 LysSerSerArgSerSerSerLeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSer 1907
Db 5941 AAATCTAGTCAAGACAGTCTCTTGGGAAATAAAGCCCCGACTTCAGGTAACTGTCT 6000
Qy 1908 GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLysHisProProGlyAsnIle 1927
Db 6001 GGTCAAGATGACGCTTCAAGTCTTGTGACCCCGACAGACCTTCCACCTCTGGCAACATC 6060
Qy 1928 ProGluSerGlyGlnAsnGlnLeuLeuGlnProLeuLysProSerProSerSerAspAsn 1947
Db 6061 CCAGAGTCCGGGAGAAATCAGCTGTACAGCCCTTAAAGCCATCTCCCTCAGTAGCAAC 6120
Qy 1948 LeuTySerAlaPheThrSerAspGlyValIleSerValProSerLeuSerAlaProGly 1967
Db 6121 CTCTATTCAGCTTACCAAGATGTGCATTTTCAGTACCAAGCCTTTCGCTCCAGT 6180
Qy 1968 GlnGlyThrSerSerThrAsnThrValGlyValAlaThrValAsnSerGlnAlaAlaGlnAla 1987
Db 6181 CAAGGACCCAGACGACCAAAACATGTTGTGGGCAACAGTGAACAGCCAGCCAGCT 6240
Qy 1988 GlnProProAlaMetThrSerSerArgLysGlyThrPheThrAspAspLeuHisLysLeu 2007
Db 6241 CAGCCTCTGCGCATGACGCTCCAGCAGGAAGGACATTCACAGATGACTTGACACAATGG 6300
Qy 2008 ValAspAsnThrPalaArgAspAlaMetAsnLeuSerGlyArgArgGlySerLysGlyHis 2027
Db 6301 GTAGACAATTTGGGCGCGAGATGCCATGATCTTCAAGCAGAGAGAGAAAGGAGAC 6360
Qy 2028 MetAsnTyArgGluGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCysIle 2047
Db 6361 ATGATTTATGAGGCGCTCGGAATGGCAAGAAATTTCTGCACCTGGGCAACTGTGCATC 6420
Qy 2048 SerMetThrSerAsnLeuGlyLysSerAlaProIleSerAlaAlaSerAlaThrSerLeu 2067
Db 6421 TCCATGACCTCGAAGCTGGGTGGCTCTGCCCTCCCATCTCTGACAGATCAGTCACTCTCTA 6480
Qy 2068 GlyHisPheThrLysSerMetCysProProGlnGlnTyArgLysPheProAlaThrProPhe 2087
Db 6481 GGTCACTTCAACCAAGTCTATGTGTGCCCCCAAGCAAGTATGTCTTCCAGCTAACCCCATTT 6540
Qy 2088 GlyAlaGlnTPSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlnPheGlnPro 2107
Db 6541 GGCGCTCAATGAGAGGAGCGGGTGGCCAGCAGCAGCCACTTGGCGCATGTTCCAACT 6600
Qy 2108 ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerLysSerAsn 2127
Db 6601 GTGGGAAGTGGCTCTCTCAGAAATTTCAACATCAGCAATTTGCAAGAAATTCATCAGCAAC 6660
Qy 2128 ProProGlySerAsnLeuArgThrThr 2136
Db 6661 CCCCAGGCTCCAACTCGGAGCACT 6687

RESULT 4
 ABZ59199
 ID ABZ59199 standard; DNA; 7149 BP.
 XX
 AC ABZ59199;
 XX
 DT 29-MAY-2003 (first entry)
 XX
 DE Human WNK1 protein encoding DNA.
 XX
 KW WNK1, WNK4; hypotensive; gene therapy; antisense therapy; vaccine;
 XX pseudohypoadosteronism type II; PHA II; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT CDS 1..7149
 FT /*tag= a
 FT /product= "WNK1"
 XX
 PN MO2003007793-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 17-JUL-2002; 2002WO-US022744.
 XX
 PR 17-JUL-2001; 2001US-0306084P.
 XX
 PA (UTTA) UNIV YALE.
 PA (LIFT/) LIFTON R P.
 PA (WILS/) WILSON F H.
 PA (CHOA/) CHOATE K.
 PA (NELS/) NELSON-WILLIAMS C.
 PA (ISHI/) ISHIKAWA K.
 XX
 PI Lifton RP, Wilson FH, Choate K, Nelson-Williams C, Ishikawa K;
 XX
 DR WPI; 2003-289755/28.
 DR P-PSDB; ABP71619.
 XX
 PT Mammalian WNK nucleic acids and proteins useful for the diagnosis,
 PT prevention and treatment of hypertension and pseudohypoadosteronism
 PT type II.
 XX
 PS Claim 3; Page 97-101; 116pp; English.
 XX
 CC The invention relates to novel nucleic acids encoding mammalian WNK,
 CC especially human WNK1 and WNK4 proteins. The WNK polypeptides can be
 CC expressed by standard recombinant methodology. The WNK polynucleotides,
 CC polypeptides and modulators may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate WNK expression such
 CC as hypertension and pseudohypoadosteronism type II (PHA II). The
 CC present sequence represents a DNA encoding the human WNK1 protein
 CC
 SQ Sequence 7149 BP; 1933 A; 1968 C; 1622 G; 1626 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 0 Length: 7149
 Score: 10677.00 Matches: 2135
 Percent Similarity: 89.63% Conservative: 0
 Best Local Similarity: 89.63% Mismatches: 1
 Query Match: 98.75% Indels: 246
 DB: 8 Gaps: 1
 US-10-010-720-14 (1-2136) x ABZ59199 (1-7149)
 QY 1 MetSerGlyGlyAlaAlaGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 DB 1 ATGCTGGCGCGCGCAGAGAGAGAGCACTCCCGTTCCTGTTCTCTCGCG 60
 QY 21 ProAlaProAlaProLysAsnGlySerSerSerValGlyGluLeuGly 40
 DB 21 |||||

DB 61 CCGGCTCTGCCCCAGAGATGGCTCCAGCTCCGATTCCTCCGTGGGGAGAAAAGTGGGA 120
 QY 41 AAlaAlaAlaAlaAspAlaValThrGlyArgThrGluGluYrArgArgArgHisThr 60
 DB 121 GCCCGCGCGCGCGAGCTGTGACCGCGAGACCGGAGAGTACAGCGCGCGCACACT 180
 QY 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrThrGluHisArgPhePhe 80
 DB 181 ATGGACAGAGACAGCGGTGGGGGCGCGCGAGCACTACCACTGACGACCGCTTCTTC 240
 QY 81 ArgArgSerValIleCyAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 DB 241 CGCCGAGACGTCATCTGCGACTGCCAATGCCACTGCACTGGAGCTTCCGGGCTTCTCT 300
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
 DB 301 TCCCTGCCCCAGCCAGCATCCCGCGCTGCTCCGAGAGTGTCCACCGAGCCGAC 360
 QY 121 ArgGluGluThrValThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 DB 361 CGGAGAGAGACCGTGACCGGCCACCTTCCAGGTAGCCGAGCACTTCCAGCCGCT 420
 QY 141 AAlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 DB 421 GCCGCCCCCTGGGGAACAGCGCGTCCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProProPro 180
 DB 481 AGCAAGACCCCGCAGGTCCAGCTTCCAGCTTGGGGAGCAAGAGAGAGCGCGCGC 540
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGlnGluAlaArgSerGlnGln 200
 DB 541 GCGAGAGTGGCAGCGCGCGCGCGCGCGCCAGCAAGAGCCAGAGAGAGAGAGAGAGAG 600
 QY 201 GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
 DB 601 CAGATGATATCGAAGAGCTGAGAGCCAGCGCGGTGGAGATGTCTAAGATGGCGCTTT 660
 QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
 DB 661 CTCAAGTTTGACATCGAATGCGAGAGCGCTCTTTAAGAGGCTTCAAGAGGCTTGAC 720
 QY 241 ThrGluThrThrValGluValAlaTyrPysGluLeuGlnAspArgLysLeuThrLysSer 260
 DB 721 ACTGAACCAACCGCGAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 261 GluArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280
 DB 781 GAGAGCGAGAGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 QY 281 ValArgPheTyrAspSerThrGluSerThrValLysGlyLysLysCysIleValLeuVal 300
 DB 841 GTTGAATTTTATGATCTCTGGGAATCCACAGTAAAGAGAGAGAGAGAGAGATGTTGTG 900
 QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysAspPheValMetLys 320
 DB 901 ACTGAACCTTATGAGCTGTGAGACCTTAAGACGATCTGAAAAGGTTTAAAGTATGAGAG 960
 QY 321 IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
 DB 961 ATCAAGGTTCTTAAGAACTGTGCGCTGAGTCTTAAAGGTTCTTCAATCTTCAATCT 1020
 QY 341 ArgThrProPoleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 DB 1021 CGAAGCTCACTATCATTCACCGGATCTTAAATGATACCAACATTTATTCACCGGCTT 1080
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysAspAlaSerPheAla 380
 DB 1081 ACTGGCTCAGTCAGATTTGAGAGACTGCTGTGCAACCTCGAGAGCGGCTTCTTTGCC 1140
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
 DB 1141 AAGAGTGTATGATGATCCCGAGAGTTCATGCGCTGAGATGTATGAGAGAAATATATAT 1200

QY 401 GluSerValAspValIleTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db 1201 GAATCCGTTGACGTTTATGCTTTGGGATGTGCATGCTTGAGATGGGCTACATCTGAATAT 1260
QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrAlaGlyValThrSerGlyValIlys 440
Db 1261 CCTTACTCGGAGTGCAGAAATGCTGCGCAGATCTACCTCGCGACCATGGGCTGTAAG 1320
QY 441 ProAlaSerPheAspIleValAlaIleProGluValIleGluGluIleGluGluIleCysIle 460
Db 1321 CCGCCAGTTTGGACAAAGTACCAATCTCTGAAGTGAAGAAATTAATTAAGAAATGACATA 1380
QY 461 ArgGlnAsnIleAspGluArgTyrSerIleIleAspLeuLeuAsnIleAlaPhePheGln 480
Db 1381 CGACAAACCAAGATGAAGAAATATTCATCAAAAGACCTTTGAACCATGCTTCTTCCAA 1440
QY 481 GluGluThrGlyValAlaGlyValGluLeuAlaGluGluAspAspGlyGluIleAlaIle 500
Db 1441 GAGGAAACAGAGATACGGGTGAATTAAGCAGAAAGATGATGAGGAAAAATAGCCATA 1500
QY 501 LysLeuThrLeuArgIleGluAspIleIleValLeuIleValIleValIleValIleValIle 520
Db 1501 AAATTATGCTACGATTAAGATTAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 1560
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 1561 GCTATTGAGTTTCTTTTGAATTTAGAGAGATGCTCCAGAAAGATGTTGCACAAAGAAATG 1620
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisIleValThrMetAlaValAlaIleAsp 560
Db 1621 GTACAGTCTGGGTATGCTGTGAAGGTGATCAAAAGCATGGCTAAAGCTATCAAAAGAC 1680
QY 561 ArgValSerLeuIleValArgValArgGluGlnArgGluValArgGluGlnGluGlu 580
Db 1681 AGAGATCATTAATTAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1740
QY 581 LysIleValGlnGluGluSerSerLeuIleValGlnIleValGlnIleSerSerAlaSerGln 600
Db 1741 AAAAAAAGCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 601 ThrGlyIleValGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1801 ACAGAAATCAAGACAGCTCCCTTCTGCTGACACCGGACATACCTACTGCTTCAACACTCA 1860
QY 621 AlaSerValSerThrGlnValGluProGluGluProGluValAspGlnHisGlnGluLeu 640
Db 1861 GCTTCAGTTCTACCAAGATGAACCTGAAGAACTGAGGACATCAACATCAACAACTA 1920
QY 641 GlnTyrGlnIleProSerIleSerValIleSerAspGlyThrValAspSerGlyGlnGly 660
Db 1921 CAGTACCAAGCAACCAAGTATATCTGTATCTGATGGGACGGTTGACAGTGGTCAAGGA 1980
QY 661 SerSerValPheThrGluSerArgValSerSerGlnIleThrValSerTyrGlySerGln 680
Db 1981 TCCTCTGCTTCCACGAATCTCGAGTGAAGCCCAACAGACAGATTCATATGTTCCCAA 2040
QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2041 CATGAACAGGCACATTTCAAGGACAGTCCAGGGGCAATATCCTTCACTGCTCAACCA 2100
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 2101 CAGTTCAGCCCATGGGATATATCCACCTCAAGTGTGGCAGGGGCAAGGCAAGGCAAGGT 2160
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2161 CAGCCATCTCAAGATGCTTAACAGGGGTTTCATCTCCAAACCATCAACATCTCTGAG 2220
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnIleThrValGlnTyrSerLeuSer 760
Db 2221 CAGCAGCAGGAAATACAGACAGACCCCTCTCAACAGACAGTGCAGATTCCTCTTCA 2280

QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2281 CAGACATCAACCTCCAGTAGAGGCGACATCTGACACAGCAAGTAGAGTCAAGCTTCA 2340
QY 781 GlnValLeuProGlnIleValSerAlaGlyIleGln 791
Db 2341 CAACTTGTGCTCAAGATACAGCTGGAAGAAACAGCTTCCAGTTTCCAGCAGTCAACT 2400
QY 791 791
Db 2401 ATCCAAAGCCAACTCAGATCCAGATCCAGTCCAGACAAACCTGGTGTTCAGTCCACTCT 2460
QY 791 791
Db 2461 GGTCCTATTTCTTCCAGTGGAGACAGCGGCTCCCTACTGCTTGTCTCCAGTCAACCT 2520
QY 791 791
Db 2521 GTCTTCAGATTTCCCATATCAACTCTCATGTGTCTACGGCTCAGACAGGTTTCTATCC 2580
QY 791 791
Db 2581 CTTCCTCATCAATGGCAGCTGCGATTACTCAGCTCTGCTCAGCTTGCTTCATCTGCT 2640
QY 791 791
Db 2641 ACAACAGCTGGATCCCGGGGTATCAACTGTGTTCTTACTAGTCACTTCAACCTTCTG 2700
QY 791 791
Db 2701 CAGCTGTGACTGCTGCTCCAAAGTCAAGTTCACCCACAGCTCTTCAACACAGAGTTGAG 2760
QY 791 791
Db 2761 TCATGGGAATACCACTAAGCTTGGACAAAGCTGTAGTTCCACTTCTCTGGAGAT 2820
QY 791 791
Db 2821 GTTCTGTACAGAGGCTTCCCACTCGACTGCCACACAGTACCAGAGATTAATAT 2880
QY 791 791
Db 2881 GCTCCCTTCCCAAGTGGCTTCTGTTCATTCATTCATTCATTCATTCATTCATTCATTC 2940
QY 791 791
Db 2941 CCGACAGAAATCTGGCTACACTGGTACTTTCCTCCACAGTGGTCAAGCTTATGTGAA 3000
QY 791 791
Db 3001 TCAATCTTTAGTCTATGAGTGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
QY 792 792
Db 3061 GGGAGTTTAGCAAGCCCCCACTACATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
QY 795 GlnValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
Db 3121 GGAATCTTCAGGTTGCTCTCGCAGAGCCAGTTCAGTACAGAGAGAGAGAGAGAGAGAG 3180
QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
Db 3181 CCGACCACTTTGGCTTCTCTGTGAGACAGTCACTTCAATGATGTTGCTTCAAGTATGAGT 3240
QY 835 AspGlyAsnGluAsnValProSerSerSerIleValIleGlnIleValIleValIleValIle 854
Db 3241 GATGGCAATGAAAGCTCCCATCTTCCAGTGAAGGCAATGAAGAAAGCAACAAAGGG 3300
QY 855 HisTyrArgLysSerValArgSerArgSerArgHisGlnIleValIleValIleValIleVal 874
Db 3301 CATTAACGAAATCTGTAGAGAGTGCCTCTGACATGAAGAAATCTTCAAGCCCAAAATTA 3360
QY 875 ArgIleLeuAsnValSerAsnIleValAspArgValValGluCysGlnLeuGluThrHis 894

```
Db 361 AGAATTGTAATGTTCAATATTAAGAGCCGAGTAGTAGATGTCATTAGAGCTCAT 3420
Qy AsnAglYMeVAlThrPheLYSPhEaSPleuAaSPGlyASpASnProGluGluIleAla 914
Db 3421 AATAGGAAAAAGGTGTACATTCAAATTTGACCTTAGATGGTGAACACCCGAGAGATAGCA 3480
Qy ThrIleMeVAlAsnAaSPheIleuAlaIleGluArgGluSerPheValASpGln 934
Db 3461 ACAATTATGATGCAACAATGACTTTATTCTAGCAATAGAGAGAGAGTCTTTGTGATGCA 3540
Qy ValAIGluIleIleGluLYSaSPGluMeLLeuSerGluASpValSerValGluPro 954
Db 3541 GTGCAGAAATATTAAGAAAAAGCTGATGAATGCTCAGTAGAGATGTCAGTGTGAGACCA 3600
Qy GluGluYASpGlnGluLYLeuGluSerLeuGlnGlyLYSaSPaSPTYGlyPheSerGlySer 974
Db 3601 GAGGAGTACAGAGGAGTGGAGAGTCTACAGAGAAAGATGACTATGGCTTTTCAGGTTC 3660
Qy GluLYSLeuGluGluLYPheLYSGLnProIleProAlaSerSerMetProGlnGlnIle 994
Db 3661 CAGAAATTTGAAAGAGAGATTCAACAACCAATTCCTGCTCTTCATGCCCACAGCAATA 3720
Qy GlyIleProThrSerSerLeuThrGlnValValHisSerIleAGLYArgArgPheIleVal 1014
Db 3721 GGCAATTCCTACCAAGTCTTTAACTCAAGTTGTTCAATCTCGCGGAAAGCGGTTATAGTG 3780
Qy SerProValProGluSerArgLeuArgGluSerLYSValPheProSerGluIleThrASP 1034
Db 3781 AGTCTGTGCCAGAAAGCCGATACGAGAAATCAAAAGTTTCCCGAGAAATTAACAAT 3840
Qy ThrValAlaAlaSerThrAlaGlnSerProGlyMeLLeuSerHisSerAlaSerSer 1054
Db 3841 ACAGTGTGCTGCTCTACAGCTCAGAGCCCTGGAAAGAACTTGTCTCACTTCGCATATCC 3900
Qy LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsn 1074
Db 3901 CTTAGCTACAAACAGGCTTTTCTGAACCTTAGACGTGCCCAATGACAGAAAGACCAAT 3960
Qy ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProPheLeu 1094
Db 3961 AOAAGACTCCAAACTTATGATACAGAGCCAAATTTCCAGTAGTACCTCTTCTTA 4020
Qy SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 4021 AGTAGCAATGTGTGAGTCCCAACACAGACAGCAGACAGACAGTCCCTGCACAAAGC 4080
Qy SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 4081 AGCCCTCTAATGACATTTCCACATCAGTAATTCAGTCTGAGGTTACAGTCCCACTGAA 4140
Qy GluGlyIleAlaGlyValAlaThrSerThrGlyValValIleThrSerGlyLYLeuProIle 1154
Db 4141 GAGGAGATGTGTGAGGTGCCACGACAGAGTGTGTACTTCAAGGTGCTCCCACTA 4200
Qy ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 4201 CCACGTGTGTGATACCAAGTACTTCCAGCGTAGTTTCAAGTATACAAATACCTGCA 4260
Qy ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 4261 GTTCTCTCAATATCTACATACATCCCGCTCACTCAAGTCCCAATCCACATCCATGAGATC 4320
Qy ValValSerSerThrAlaLeuThrProSerValThrValSerAlaIleThrSerAlaSerIle 1214
Db 4321 GTTGTTCCTATGACGACGCTGATCTCTCAGTAACAGTTTCAAGCAACTTCAAGCTCTGCA 4380
Qy GlyGlySerThrAlaIleThrProGlyProLYSProProAlaValValSerGlnAlaAla 1234
Db 4381 GGGGGCAGTACTGTGCAACCCAGGTCTTAAGCTCCAGCTTAGATATCTCAGACAGGACCA 4440
Qy GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 4441 GGCAGCACTACTGTGAGAGCCACATTAATCAGTTTCTACACCACTTCATTTCCCAAGC 4500
Qy ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrLeuAlaGlu 1274
Db 4501 ACAGCTTACAGCTGTCCATTCAGCTTACAGACAGTACTTCTTACTTCACTTACTGAA 4560
Qy ThrValValValSerAlaHisSerLeuAspLYSThrSerHisSerSerThrThrGlyLeu 1294
Db 4561 ACCGTGTAGTTAGGCGCACACTACTAGATTAAGATCATCATTAGCAAGTACAACTGATTTG 4620
Qy AlaPheSerLeuSerAlaProSerSerSerSerProGlyAlaGlyValSerSerTYr 1314
Db 4621 GCTTCTCTCCTCTCGACCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4680
Qy IleSerGlnProGlyLYLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4681 ATTCTCAGCTGGGTGGCTCATCTCTTGGTCATTCATCATGATAGCTTCTTCTCT 4740
Qy IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 4741 ATTCTTCCCAAGCAGCAGACCTTCTTACACCTTATTAACCCCAAGTACATGATATC 4800
Qy ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4801 CCACCTTGTATACAGCTGTGGCCAAATGCTGCTGTATAGCAGACACTAATTCATAGT 4860
Qy GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer 1394
Db 4861 CAGCTCAACAGCTTGTGCTTCCCAACCAAGCCCATATCTCATTTGCTCGAAGTATCTCT 4920
Qy AspThrGlnProLYSAlaProGlyIleAspAspIleLYSThrLeuGluGluLYLeuArg 1414
Db 4921 GATACACACCCCAAGCTCCCTGGAAATTGATACATTAAGAACTTCTAGAAAGAAAGCTGGG 4980
Qy SerLeuPheSerGluHisSerSerSerSerGlyValaGlnHisAlaSerValSerLeuGluThr 1434
Db 4981 TCTCTGTGATGAAACAGCTCATCTGAGCTCAGCATGCTGCTGTCTCACTGGAAGCC 5040
Qy SerLeuValIleGluSerThrValThrProGlyIleProThrThrAlaAlaProSer 1454
Db 5041 TCACCTAGCATAGAGAGACGCTCACACAGCAATCCCACTGCTGCTGACCAAGC 5100
Qy LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 5101 AAACCTCTGACTTCTACCAAGATCTTGTCTTACCAACCAATTTCCACTGAGAACCA 5160
Qy ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 5161 GTTCTTGTGCAATTAACCAAGTGTGTCACACTGGGCAAGTTTCTACCCCAAGTACGACT 5220
Qy ThrThrSerGlyValLYSProGlyThrAlaProSerLYSProProLeuThrLYSAlaPro 1514
Db 5221 ACTACATCAGAGAGGAAACCTGGAATGCTCTCCCAAGCCACTTAACTAAGCTTCG 5280
Qy ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGluLeuProPro 1534
Db 5281 GTGCTGCAAGGGTACTGAATCTTCCAGCAGTACTTACCAACAGCAGAGCTGCACCT 5340
Qy PheProGlyProSerLeuThrGlnSerGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 5341 TTTCAGAGACCTTCTTAACCCAGTCCAGCAACCTTCAAGAGATCTTGATGCTCAATTG 5400
Qy ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValAlaGlyProValSerMet 1574
Db 5401 AGAAGAACACTTATCTCAGAGATGATACATGACTTCTGTGGTGTGCTGTGTTCATG 5460
Qy AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnLYSGLYValSerGlnVal 1594
Db 5461 GGCGCTCAACAGCAATCAACAGAGCAGAAACACAGCTCAGAGAGGTGTTTCAAGTC 5520
Qy LysGlnGlyProValLeuAlaIleThrSerSerGlyValaGlyValPheLYSMeGlyArgPhe 1614
Db 5521 AAAGAAAGCCCTGTCTTAGCAACTAGTTCAAGAGACTGGTGTGTTTAAAGATGGAGATTT 5580
```

QY 1615 GlnValSerValAlaIaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
 Db 5581 CAGGTTTCTGTGGACGACGAGCGTGCCGAGAAAGGGGTAAATAATCTCAGAAAGATCA 5640
 QY 1635 LysSerValHisPheGlnSerSerThiSerGlnSerValLeuSerSerSerPro 1654
 Db 5641 AAGCTGTTCATTTTGAACTCAGACCTCAGAGTCTCCAGTCTATCATCAAGTAGAGTCA 5700
 QY 1655 GlnSerThiLeuValLysProGlnProAsnGlyIleThriIleProGlyIleSerSerAsp 1674
 Db 5701 GAGGATACCTTGTTGAAACAGAGCCGAGTGCATACCACTCCCTGATCTCTTCAGAT 5760
 QY 1675 ValProGlnSerAlaHisLysThiThiAlaSerGlnAlaLysSerAspThiGlyLanPro 1694
 Db 5761 GTGCCAGAGAGTCCCAAAACTCTGCTCAGAGGCAAACTCAGACACTGGGCAAGCT 5820
 QY 1695 ThrLysValGlyArgPheGlnValThiThiThiAlaAsnLysValGlyArgPheSerVal 1714
 Db 5821 ACCAAGTTGGACGTTTCAAGTGCACACTACAGCAAAAGTGGGTCTCTCTGTA 5880
 QY 1715 SerLysThiGlnAspLysIleThiThiAspThiLysLysGlnGlyProValAlaSerProPro 1734
 Db 5881 TCATAAAGCTGAGAGACATCATGACACAAAGAAAGAGACAGTGGCACTCTCTCTCT 5940
 QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlnLysProGln 1754
 Db 5941 TTTATGATTTGGAAACAGAGCTGTTCTTCTGCTGATACCAAGAAAGAGAGAGCTGAA 6000
 QY 1755 LeuSerGlnProSerHisLeuAsnGlyProSerSerAspProGlnAlaIlePheLeuSer 1774
 Db 6001 CTGTGAGAGCTTCATCATCTAAATGGCGCTCTTCTGACCGGAGGCGCTTTTAAAGT 6060
 QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
 Db 6061 AGGGAATGGATGATGGTTCGGGTAGTCCACACTGCGCCCACTACAGCTCAAGAGGC 6120
 QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSer 1814
 Db 6121 CTTCCTACCCAGAACTCAAGTCAAGCTTGTATTAATTAATTAATTAATTAATTAATTA 6180
 QY 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuGlnLeuArgArgLeuArg 1834
 Db 6181 AGCGACATGAGTCAAGATATGAGATGAGAGCTTAAAGTTAAGAGCTCGACGACTGACA 6240
 QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
 Db 6241 GATTAACATCTCAAGAGATTCAGAGCTGCGAGAGTCCCGAGAACATGAATGAATCT 6300
 QY 1855 LeuTyThiLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu 1874
 Db 6301 TTGTATACCAAACTGGGCAAGGTGCCCCCTGTGTATTAATTAATTAATTAATTAATTA 6360
 QY 1875 SerGlyArgArgArgArgProThiLysSerLysGlySerLysSerSerSerSerSer 1894
 Db 6361 TCAGGGAGAGAGACGACGACCACTAAAGCAAAAGGACCAAAATCTAGTCAAGAGTTC 6420
 QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyLysLeuSerGlyGlnSerAlaIleSerVal 1914
 Db 6421 TTGGGGAAATAAAGCCCGAGCTTCAGGTAACTGCTGAGCAAGTGCAGGCTTCAGTTC 6480
 QY 1915 LeuHisProGlnGlnThiLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
 Db 6481 TTGCACCCCGACAGACCTCCACCTCTGCGAACATCCAGAGTCCGGGAGAGATGAG 6540
 QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuTySerAlaPheThiSer 1954
 Db 6541 CTGTATACAGCCCTTAAGCACTCTCCCTCAGTGCACCACTCTATTCAGCTTCACCACT 6600
 QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThiSerSerThiAsn 1974
 Db 6601 GATGATGCCATTTCTAGTACCAAGCTTTCTGCTCCAGGTCAAGGAGAACGACGACCAAC 6660

QY 1975 ThrValGlyAlaThiValAsnSerGlnAlaIaGlnAlaGlnProProAlaMetThiSer 1994
 Db 6661 ACTGTTGGGGCAACAGTGAACAGCCAGCCGCAAGCTCAGCTCTGACATGACGTCC 6720
 QY 1995 SerArgLysGlyThiThiPheThiAspAspLeuHisLysLeuValAspAsnTTPAlaArgAsp 2014
 Db 6721 AGCAGGAAGGGCAATTCACAGATGACTTGGACAAAGTTGGTGAACAAATGGGCGCGAGAT 6780
 QY 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyGlnGlyProGly 2034
 Db 6781 GCCATGAATCTCTCAGGACGAGAGAGAGAACCAAGGCAATGAATTAAGAGGGCCCTTGA 6840
 QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThiSerAsnLeuGly 2054
 Db 6841 ATGGCAAGAAAGTCTCTGACACTGGGGAATGTGATGATCTCATAGACTTCGACCTGGCT 6900
 QY 2055 GlySerAlaProIleSerAlaIleSerAlaThiSerLeuGlyHisPheThiLysSerMet 2074
 Db 6901 GGCCTGCGCCCACTCTGACGATCAGCTACCTCTTACGTCACTTCACCAAGTCTATG 6960
 QY 2075 CysProProGlnGlnTyGlyPheProAlaThiProPheGlyAlaGlnTTPSerGlyThi 2094
 Db 6961 TGCCCCCAGACGAGTATGGCTTTCCAGCTAACCCCATTTGGCGCTCAATGAGTGGACG 7020
 QY 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThiAlaSerLeuGln 2114
 Db 7021 GGTGGCCAGACACCAAGCCACTTGGCCAGTTCCAACTGTGGGAAGTGGCTCTTGAG 7080
 QY 2115 AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg 2134
 Db 7081 AATTTCACACATCAGCAATTTGCAAGAAATCCATCAGCAACCCCGAGGCTCCAACTGGG 7140
 QY 2135 ThrThi 2136
 Db 7141 ACCACT 7146

RESULT 5
 ACH03764
 ID ACH03764 standard; DNA; 7149 BP.
 XX
 AC ACH03764;
 XX
 DT 26-SEP-2003 (first entry)
 XX
 DE DNA encoding novel human protein kinase #1.
 XX
 KM Human; kinase; ds; gene; cosmetic application; nutraceutical application.
 XX
 OS Homo sapiens.
 XX
 PN US6541252-B1.
 XX
 PD 01-APR-2003.
 XX
 PF 14-MAY-2001; 2001US-00854856.
 XX
 PR 19-MAY-2000; 2000US-0206015P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hilbun E, Donoho G, Turner CA;
 XX
 DR MPI; 2003-575927/54.
 DR P-PsDB; ABO44381.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 XX nutraceutical applications.
 PS Discloure; Page; 11pp; English.
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents DNA encoding

Db 1921 CAGACACGACGACCGATATATCTGTATCTGATGAGGAGCGTTGACAGTGTCCAGGGA 1980
 QY SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 Db 1981 TCCCTGCTTCCACGAAATCTCGAGTGGAGCGACCAAGACAGTTTCAATAGTGTCCAA 2040
 QY HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 Db 2041 CATGACAGCGCACATTTCTACAGGACAGTCCAGGGGATATACCTTCTACTGTCCAGCA 2100
 QY GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 Db 2101 CAGTCTAGCCCCATGGGATATCTCACTCAAGTGTGGCAGGGGCAAGCGCAGCGGT 2160
 QY GlnProSerSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
 Db 2161 CAGCCATCTCTAAGTAGCTTACAGGGGTTTCATCTTCCAAACCATCAACATCTTCAG 2220
 QY GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
 Db 2221 CAGCAGCAGGGAATACAGCAGACAGCCCTCTCTCAAGCAGACAGTGGCATTTCACTTCA 2280
 QY GlnThrSerThrSerSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 Db 2281 CAGCATCAACCTCCAGTGGAGCCACTACGACAGCCAGTGTGAGTCAAGCTCAAGCTCCA 2340
 QY GlnValLeuProGlnValSerValSerAlaGlyValGln----- 791
 Db 2341 CAAGTCTTGCTCTCAAGTATCAAGTGTGAAAACAGCTTCCAGTTCCCAAGCCAGTCAACT 2400
 QY ----- 791
 Db 2401 ATCCAGGCGAAGCTCAGATCCAGTTGGCAGACACAAACCTCGGTTGTTCCAGTCCACT 2460
 QY ----- 791
 Db 2461 GGTGCTCATTTCTCTCAGTGGGACAGCCGCTCCCTACTCCCTTGCTCCCTCAGTACCT 2520
 QY ----- 791
 Db 2521 GTCCTCAGATTCCTCATATCACTCTCATGTGTCTACGGCTCAGACAGGTTTCTCATCC 2580
 QY ----- 791
 Db 2581 CTTCCCATCACAATGGAGCGTGGCAATTACTCAGCCTGTGCTCAGTTGGCTTCACTGCT 2640
 QY ----- 791
 Db 2641 ACAACAGCTGCGATCCCGGGGGTATCAACTGTGTTCTTACTAGCTTCCAACCTTCTG 2700
 QY ----- 791
 Db 2701 CAGCCTGTGACTCAGCTGCCAAGTCAAGTTCACCCACAGCTCTTCAACACAGCACTTCA 2760
 QY ----- 791
 Db 2761 TCCATGGGAATACAGAGTAACCTTGGACAGACGTGTAGATTCCACTTCTCTGGAGAT 2820
 QY ----- 791
 Db 2821 GTTCTGTACAGAGGCTTCCACTCGACTGCCACACAGTACCCAGAGATTCAAAATTT 2880
 QY ----- 791
 Db 2881 GCTCCCTCTTCCAAAGTGGCTTGTGTTGATTCATTTCTACAGTCTTAACTCCCTCCATG 2940
 QY ----- 791
 Db 2941 CCGACAGAGTACTGGCTACACCTGGGTACTTTCACAGTGGTGGAGCCTTATGTGAA 3000
 QY ----- 791

Db 3001 TCAATCTTTAGTTCCTATGGGTGTGTAGAGACAGGTTCAAGTCTCCACACGGA 3060
 QY ----- SerThrGln 794
 Db 3061 GGGAGTTAGCAGACGCCCCCATCAATCTCTCCACAGACAGGATTTGGAGAGTACTAG 3120
 QY GlnValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
 Db 3121 GGAATCTCTAGAGTTGCTCTCGACAGACAGTTGCAGTACAGACAGCCCAAGTACCCAG 3180
 QY ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 Db 3181 CCGACCACTTGGCTTCTCTGTAGACAGTGCACATTCAGATGTGTGCTTCAGTATGAGT 3240
 QY AspGlyAsnGlnValAsnValProSerSerSerGlyValArgHisGlyValGlyThrThrValArg 854
 Db 3241 GATGGCAATGAAAGCGTCCATCTTCCAGTGAAGGCGATGAAGAGAACTACAAACGG 3300
 QY HisTyrArgLysSerValArgSerArgSerArgHisGlnLysThrSerArgProLysLeu 874
 Db 3301 CATTACCGAAATCTGTAAAGAGTGGCTCTGACATGAAAACTTCACGCCCAAAATTA 3360
 QY ArgIleLeuAsnValSerAsnLysGlyAspArgValValGlyCysGlnLeuGlnThrHis 894
 Db 3361 AGAATTTGATGTTTCAATTAAGAGACCGAGTATGATGATCAATTAGACATCAT 3420
 QY AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAla 914
 Db 3421 AATGGAATATGTTCAATTCATTAATTTGACCTTAATGTGTACAAACCCGAGAGATACCA 3480
 QY ThrIleMetValAsnAsnAspPheIleLeuAlaIleGlnArgGlnSerPheValAspGln 934
 Db 3481 ACAATTATGTGAAACATGACTTATTCTGCAATAGAGAGAGTGGTTGTGTGATCAA 3540
 QY ValArgGlnIleIleGlnLysAlaAspGlnMetLeuSerGluAspValSerValGluPro 954
 Db 3541 GTGGAGAAATTTATGAAAAAGCTGATGAATGTCTGAGTGTGATGTGATGAACCA 3600
 QY GlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
 Db 3601 GAGGATATCAGGATTTGAGAGTCTTACAAAGAAAGATGACTATGGCTTTTCAGGTCT 3660
 QY GlnLysLeuGlnGlyGlnPheLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
 Db 3661 CAGAAATTTGAGAGAGGTTCAACCAACCAATTCCTCGCTTCCAGGCCACACAAATA 3720
 QY GlyIleProThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgPheIleVal 1014
 Db 3721 GGCATTTCTTACCACTTTTAACTCAAGTGTTCATCTGGGGAAGCGGTTTATAGTG 3780
 QY SerProValProGlnSerArgLeuArgLysSerLysValPheProSerGlnIleThrAsp 1034
 Db 3781 AGTCTTGTCCAGAAAGCCGATTCAGAAATCAAAAGTTTCCCAAGTAAATACAGAT 3840
 QY ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
 Db 3841 ACAATGTGCTGCTTACAGCTCAGAGCCCTGGAATGAACCTTCTCACTGCACTCATATCC 3900
 QY LeuSerLeuGlnGlnAlaPheSerGlnLeuArgArgAlaGlnMetThrGlnGlyProAsn 1074
 Db 3901 CTTAGTCTACAAAGGCTTTTTCGAACCTTAGAGTGTCCCAAAAGACAGAAAGGCCAAAY 3960
 QY ThrAlaProProAsnPheSerHisThrGlyProThrPheProValAlaProProPheLeu 1094
 Db 3961 ACAGCACTTCCAACTTTAGTACATACAGACCAACATTTCCAGTAGTACCTCTTTCTTA 4020
 QY SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
 Db 4021 AGTAGCATTTGCTGAGTCCCAACACAGCAGCAACCAAGACCAAGTCCCTGCAACAGC 4080
 QY SerProProAsnAspIleSerThrSerValIleGlnSerGlnValThrValProThrGln 1134
 Db 4081 AGCCTCTTATGACATTTCCATCATCATGATTAATTCAGTGAAGTTTACAGTGCACACTGAA 4140

QY	1135	GIuGIyLIleAlaGIyValAlaThnSerThrgLyValValThnSerGIyLyLeuProIle	1154
Db	4141	GAGGGAGATGCTGGAGGTGGCCACCGACAGAGTGGTGAACCTTACAGGGTCTCCCATTA	4200
QY	1155	ProProValSerGIuSerProValLeuSerSerValValSerSerIleTheIleProAla	1174
Db	4201	CCACTGCTGCTGAATCACCACTACTTCCAGCGTAAAGTTTCAAGATATACAAATCTGTCA	4260
QY	1175	ValValSerIleSerThnThnSerProSerLeuGIuValProThnSerThnSerGIuIle	1194
Db	4261	GTGGCTCAAAATCTACTACAAATCCCGCTCACTTCAAGTCCCAATCCACATCTGAGATC	4320
QY	1195	ValValSerSerThnAlaLeuTyProSerValThnValSerAlaThnSerAlaSerAla	1214
Db	4321	GTTGTTCTTAGACAGCACTGATTCCTTCAGTAAACAGTTTGAAGCAACTTCAAGCTCTGTCA	4380
QY	1215	GIyGISeSerThnAlaThnProGIyProIlyProProAlaValValSerGIuAlaAla	1234
Db	4381	GGGGGACATCTGCTACCCCAAGTCTTAAGCTCCAGCTGTAAGTATCTCAGACAGGACGA	4440
QY	1235	GIySerThnThnValGIyAlaThnIleuThnSerValSerThnThnThnSerPheProSer	1254
Db	4441	GGCAGCACTACTGTTGGAGCCACATTACATCAGTTTCTTACACACACTTCATCCCAAGC	4500
QY	1255	ThnAlaSerGIuLeuSerIleGIuLeuSerSerSerThnSerThnProThnLeuAlaGIu	1274
Db	4501	ACAGCTTACAGCTGCTCCATTCAGTTTGAAGCAGCACTACTTCTTACTCTTCACTTGAAGAA	4560
QY	1275	ThnValValValSerAlaIleSerIleuAspIlyThnSerIleSerSerThnThnGIyLeu	1294
Db	4561	ACCGTGGATGTAAGGCACACTCATATATAAAGCATCTCATAGCAGTACAACTGGATG	4620
QY	1295	AlaPheSerLeuSerAlaProSerSerSerSerSerProGIyAlaGIyValSerSerTyx	1314
Db	4621	GCTTCTCTCCCTCTGACACCAATCTCTCTCTCTCTCTCTCTCTGAGAGAGAGGTCTAAGTAT	4680
QY	1315	IleSerGIuProGIyGIyLeuIleHisProLeuValIleProSerValIleAlaSerThnPro	1334
Db	4681	ATTTCACGCTGGGGGTGATGCTCTTGGTCACTTCATCAGTAAAGCTTCTTACTCTCT	4740
QY	1335	IleLeuProGIuAlaAlaGIyProThnSerThnProLeuLeuProGIuValProSerIle	1354
Db	4741	ATTCTCCCAAGACAGACAGCAACTTCTTACACCTTATTAACCCCAAGTACCTAGTATC	4800
QY	1355	ProProLeuValGIuProValAlaAsnValProAlaValGIuGIuThnLeuIleHisSer	1374
Db	4801	CCACCTCTGTACAGCTCTGTTGCCAATGTGCTGCTGTACAGACAGACACTAATCATATGT	4860
QY	1375	GIuProGIuProAlaLeuLeuProAsnGIuProHisThnIleHisTyProGIuValAspSer	1394
Db	4861	CAGCTTCAACAGAGCTTGTGCTTCCCAACAGGCCCAACTACTATGTCTCGAAGTGAATTC	4920
QY	1395	AspThnGIuProIlyAlaProGIyIleAspAspIleTyThnIleGIuGIuLeuAlaGIy	1414
Db	4921	GATACACACACCAAAAGCTCTTGGAATTGATGACATTAAGAACTTACAGAAGAAACTGGG	4980
QY	1415	SerLeuPheSerGIuHisSerSerSerGIyAlaGIuHisAlaSerValSerLeuGIuThn	1434
Db	4981	TCTGTGTCAATGAACACAGCTCATCTGGAAGCTCAGCAATGCTGTCTCACTGAGAGCC	5040
QY	1435	SerLeuValIleGIuSerThnValThnProGIyIleProThnThnAlaValAlaProSer	1454
Db	5041	TCACTAGTCATAGAGAGCACTGTACACACAGGACATCCCACTACTGCTGTGGCCCAAGC	5100
QY	1455	LySLeuLeuThnSerThnThnSerThnTySLeuProProThnAsnLeuProLeuGIyThn	1474
Db	5101	AAAGTCCGACTTCTTACCAAAAGTACTTGCTTACACCAACCAATTATCCACTGGAACA	5160
QY	1475	ValAlaLeuProValThnProValValThnProGIyGIuValSerThnProValSerThn	1494
Db	5161	GTTCCTTGGCAGTTCACCAAGTGTGTCAACCTGGGCAAAAGTTCTTACCCCACTGAGACT	5220

QY	1455	ThrThrSerGlyValIysProGlyThrAlaProSerIysProProLeuThrTrpAlaPro	1514
Db	5221	ACTACATCAGGAGGAAACCTGGAAACGTCCCTCCCAAGCCACTCTTAACATGAAGCTCCG	5280
QY	1515	ValIeuProValIGlyThrGluLeuProAlaGlyThrIleuProSerGluGlnIleuProPro	1534
Db	5281	GTGCTGCCAGTGGGAGTGAACCTTCCAGAGGTACTCTAACCCAGCCGAGGCTCCACCT	5340
QY	1535	PheProGlyProSerIleuThrGlnSerGlnIleProLeuGluAspIleuAspIleuLeu	1554
Db	5341	TTTCCAGACCTCTCTTAACCCACTCCAGAACCTCTTAAGAGATCTTGAAGCTCAATTG	5400
QY	1555	ArgArgThrIleuSerProGlu**IleThrValThrSerAlaValIGlyProValSerMet	1574
Db	5401	AGAGAAACACTTAATCCAGAGATATATACAGTACTCTGCGGTGTGCTGTTCATG	5460
QY	1575	AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnIysGlyValSerGlnVal	1594
Db	5461	GCGGTCTCCAAACGAAATCACAGAAACGAGAAACACAGCTCCAGAAAGGGGTCTTCCAAATC	5520
QY	1595	LysGluGlyProValIleuAlaThrSerSerGlyAlaGlyValPheIysMetGlyArgPhe	1614
Db	5521	AAAGAGGCCCTGCTCTTGCACACTAGTTCAGAGACTGTGTTTTTAAGATGGACGATTT	5580
QY	1615	GlnValSerValAlaAlaAspGlyAlaGlnIysGluGlyIysAsnIysSerGluAspAla	1634
Db	5581	CAGGTTCTGTTGCAGCAGACGGTGCACAAAGAGGGTAAATAATGATCAGAGATGCA	5640
QY	1635	LysSerValHisPheGluSerSerThrSerGluSerValIleuSerSerSerPro	1654
Db	5641	AAGTCTGTCATTTTTGATCCAGCACTCAGAGCTCTCAGGCTATCAAGTAGAGTCCA	5700
QY	1655	GluSerThrIleuValIysProGluProAsnGlyIleThrIleProGlyIleSerSerAsp	1674
Db	5701	GAGAGTACCTTGATGGAACCAACGACCGAATGGCATTAACATCCCTGATATCTCTTCAGAT	5760
QY	1675	ValProGluSerAlaHisIlyThrThrAlaSerGluAlaIysSerAspThrGlyGlnPro	1694
Db	5761	GTGCACAGAGATGCCCAAAACACTGCTCCAGAGCAAAAGTCAACACTCGGGCACTCT	5820
QY	1695	ThrIysValIGlyA-GPheGlnValThrThrThrAlaAsnIysValIGlyA-GPheSerVal	1714
Db	5821	ACCAAGGTGGACCTTTTCAGGTACCAACTACGAAACAAAGTGGTCTTTCTCTGTA	5880
QY	1715	SerIysThrGluAspIysIleThrAspThrIysIysGluGlyProValAlaSerProPro	1734
Db	5881	TCAAAAACCTGAGCAACAGATCATCGACACAAAGAAAGAGGACAGTGCATCTCCCT	5940
QY	1735	PheMetAspLeuGluGlnAlaValIleuProAlaAlaIleProIysGlyIysProGlu	1754
Db	5941	TTTATGATTTGGAAACAAGCTGTTCTTCCTGCTGTATACCAAGAAAGAAACCTGAA	6000
QY	1755	IleuSerGluProSerThrIleuAsnGlyProSerSerAspProGluAlaIlePheIleuSer	1774
Db	6001	CTGTCAAGGCTTCAACATCTAATGGGCGGCTTCTGACCCGAGGCGCGCTTTTAAAT	6060
QY	1775	ArgAspValAspAspGlySerGlySerProHisSerProHisGlnIleuSerSerIysSer	1794
Db	6061	AGGATATGGATGATGTTCCGATGATGCCACACTGCCATCAAGCTGAGACTCAAAAGC	6120
QY	1795	IleuProSerGlnIleuIleuSerGlnSerIleuSerAsnSerPheAsnSerSerIysMetSer	1814
Db	6121	CTTCTTGAACCAATCTTAAGTCAAAAGCTTAAGTATTCATTTTAATCTCTCTTACATGAGT	6180
QY	1815	SerAspAsnGluSerAspIleGluAspGluAspIleuIysLeuGluIleuArgArgLeuArg	1834
Db	6181	AGCACAATATGATCAGATATCGAAGATGAAGAACTTAAGTTAGAGCTGCACACACTAACA	6240
QY	1835	AspIysHisIleuIysGlnIleGlnAspLeuGlnSerArgGlnIysHisGlnIleGluSer	1854
Db	6241	GATTAACATCTCAAAAGATTCACAGACTCGTAGAGTGCACAGAACATGAAATTTGAATCT	6300
QY	1855	IleuIysThrIysLeuGlyIysValProProAlaValIleIleProProAlaIleProIleu	1874


```
Db 481 AGCAAGACCGCCAGTCTCCAGCCTTGTGGGAGCAAGAGACCGCCGCG 540
Qy 181 ALaargserglyserglyglyseralaysgluproglingluarsergln 200
Db 541 GCGAAGATGGGAGCGCGCGCGCCAGCGCAAGAGCAAGAGAAAGCGAGCGAG 600
Qy 201 GlnasparillegluugluugluuThrlysaValaGlymetSerAspAspGlyVargphe 220
Db 601 CAGAGATGATCGAAGAGCTGGAGACCAAGCGCGGAGATGCTTAACGATGCGCTT 660
Qy 221 LeuylsPheaspillegluuilegIyarglyserPheylsThrValTyrlsGlyLeuasp 240
Db 661 CTCAAGTTTGAACATGGAATGGGACAGAGCTCTTTAAGACGCTTACAAAGCTTGAC 720
Qy 241 ThrGluThrThrValGluValAlaTrpCysgluLeuGlnasparGlyLeuThrlyser 260
Db 721 ACTGAACACCGGTGGAGTGGCTGTGTGTGAATCGAAGATCGAATAAATAAGTCT 780
Qy 261 GluArgGlnArgPheylsGluGluValaGluMetLeuylsGlyLeuGlnHisProasnle 280
Db 781 GAGAGCGAAGATTTAAAGAAAGACTGAATGTTAAAGGCTTACACATCCCAATAT 840
Qy 281 ValArgPheTyAspSerTrpGluSerThrVallysglylysCysileValleuVal 300
Db 841 GTTAGATTTTATGATTCCTGGGAATCCACAGTAAAGAAAGATGATTTTGTG 900
Qy 301 ThrGluMetThrSerGlyThrleuylsThrTyrlsValArgPheylsValMetlys 320
Db 901 ACTGAACCTTAAGACCTCTGGAAACCTTAAACGTAATCGAAGAGTTTAAAGTGTAG 960
Qy 321 IlelyValleuArgSerTrpCysArgGlnIleleuylsGlyLeuGlnPheleuHisThr 340
Db 961 ATCAAGTTCTTAAGAGCTGGTGGCTGAGATCCTTAAAGGCTTCAAGTTCTTCAATCT 1020
Qy 341 ArgThrProIleIleHisArgAspLeuylsCysAspAsnIlePheIleThrGlyPro 360
Db 1021 CGAATCCACCTATCATTCACCGGATCTTAATGTGACAAATCTTTATCACCGGCT 1080
Qy 361 ThrGlySerValylsIleGlyAspLeuGlyLeuAlaThrleuylsArgAlaSerPheAla 380
Db 1081 ACTGGCTAGCTCAAGATTTGAGAGACTCGGTCTGGGACCTGAAACCGGGCTCTTTTCC 1140
Qy 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrlsGluGlyAsp 400
Db 1141 AAGAGTGTGATGATGATCCCAAGATTCATGGCCCTGAGATGTATGAGAGAAATATAT 1200
Qy 401 GluSerValAspValTyrlaPheGlyMetCysMetLeuGluMetAlaThrSerGlyTrp 420
Db 1201 GAATCCGTTGAGTTTATGCTTTGGATGTGCATGCTTGAGATGCTAATCTGAATAT 1260
Qy 421 ProTyrlsSerGlyCysGlnAsnAlaAlaGlnIleTyrlaArgValaThrSerGlyVallys 440
Db 1261 CCTTAATCTGGAGTGGCAAAATGCTGCACAGATCTACCTCGGTGACCAAGGGGTGAG 1320
Qy 441 ProAlaSerPheAspLyValAlaIleProGluValylsGluIleIleGluGlyCysile 460
Db 1321 CCAGCGATTTTGAACAAAGTACCAATCTCTGAAGGAAAGAAATATTGAAGAGATGCTA 1380
Qy 461 ArgGlnAsnlysaSpGluArgTyrlserIlelysaPheLeuAsnHisAlaPhePheGln 480
Db 1381 CGAAGAAACAAAGATGAAGATATCTCCATCAAAAGACTTTGAACATGCTTCTTCAA 1440
Qy 481 GlnGluThrThrValArgValaGluLeuAlaGluGluAspAspGlyGluIleAlaIle 500
Db 1441 GAGAGAAACAGAGTACGGGTGAATATGCAAGAAAGATGATGAGAGAAATAATGCCCTA 1500
Qy 501 LysLeuTrpLeuArgIleGluAspIlelysaPheLeuylsGlyTyrlsAspAsnGlu 520
Db 1501 AAATTTATGGCTACGATTTGAAGATTTTAAAGAAATTTAAAGGAAATACAAAGATATAA 1560
Qy 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValaIleGlnIleMet 540

Db 1561 GCTATGAGTTTCTTTTATTATGAGAGAGATGCTCCAGAAAGATGTTGCACAAAGATNG 1620
Qy 541 ValGluSerGlyTyrlValCysGluGlyAspHislysaThrMetAlaIleAlaIlelysaSp 560
Db 1621 GTAGAGTCTGGGTATGCTGTGAAGGTGATCAAGAACCAATGGCTTAAAGCTATCAAAAG 1680
Qy 561 ArgValSerLeuIlelysaArglysaArgGluGlnArgGlnleuValaArgGluGluGln 580
Db 1681 AGAGTATATTAAATTAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Qy 581 LyslysaGlnGluGluSerSerLeuylsGlnValaGluGlnSerSerAlaSerGln 600
Db 1741 AAAAAAGCAGAGAGAGAGAGAGAGTCTCAACAGAGAGAGAGAGATCCAGTCTTCCAG 1800
Qy 601 ThrGlyIlelysaGlnleuProSerAlaSerThrGlyIleProThrAlaSerThrSer 620
Db 1801 ACAGAAATCAAGACAGCTCCCTTGTGTGACAGCGCAATACCTACTGCTTACACCTTCA 1860
Qy 621 AlaSerValSerThrGlnValaGluProGluGluProGluAlaAspGlnHisGlnGlnleu 640
Db 1861 GCTTCAGTTTCAACAGATGAACTGAAGACCTGAGCGAGATCAACATCAACACACTA 1920
Qy 641 GlnTyrglnGlnProSerIleSerValleuSerAspGlyThrValaAspSerGlyGlnGly 660
Db 1921 CAGTACAGCAACCCAGATATATCTGTATTCTGATGGACGTTGACAGTGTGACAGGGA 1980
Qy 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrlsSerGln 680
Db 1981 TCCTGTCTTCAAGATCTCGAGTGGAGAGCAAGACAGAGATTCATATGTTCCCA 2040
Qy 681 HisGluGlnAlaHisSerThrGlyThrValaProGluHisIleProSerThrValaGlnAla 700
Db 2041 CATGAACAGGACATTTCTACAGGACAGTCCCAAGGACATATACCTTCTACTGTCAGACA 2100
Qy 701 GlnSerGlnProHisGlyValaTyrlProProSerSerValaIleGlnGlyGlnSerGlnGly 720
Db 2101 CAGTCTAGCCCCAGGGGTATATCCACCTCAAGTGG----- 2139
Qy 721 GlnProSerSerSerleuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2139 ----- 2139
Qy 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValaGlnTyrlSerleuSer 760
Db 2140 ---CAGCAGGAAATACAGACACAGACCCCTCTCAACAGAGAGTGTGACTTCA 2196
Qy 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValaSerGlnProGlnAlaPro 780
Db 2197 CAGACATCAACCTCAAGTGAAGGCCACTATGSCACAGCAGAGTACAGCTCAAGCTCCA 2256
Qy 781 GlnValleuProGlnValaSerAlaGlylysaGlnSerThrGlnGlyValaSerGlnValaAla 800
Db 2257 CAAGCTTGTCTCAAGTATCAGCTGAGAAACAGAGTATCTGAGGAGTCTCAGGTTCT 2316
Qy 801 ProAlaGluProValaAlaValaIleGlnProGlnAlaThrGlnProThrThrleuAlaSer 820
Db 2317 CCTGACAGACAGTGTGAGTGCACAGCCCAACTACCCAGCCAGCACACTTGGCTTCC 2376
Qy 821 SerValaAspSerAlaHisSerAspValaIleAserGlyMetSerAspGlyAsnGluAsnVal 840
Db 2377 TCTGTAGACAGTGCATTCAGATGTTGCTTCAAGTATGATGTATGGCAATGAGAACCTGC 2436
Qy 841 ProSerSerSerGlyArgGlnHisGluGlyValaArgThrlysaArgHisTyrlArglySerVal 860
Db 2437 CCATCTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2496
Qy 861 ArgSerArgSerArgHisGlnTyrlsThrSerArgProlysaLeuArgIleleuAsnValSer 880
Db 2497 AGAGTGGCTCTGCACATGAGAAATCTTCAAGCCCAATATTAAGATTTGAATGTTTCA 2556
Qy 881 AsnlysaGlyAspArgValaValaGluCysGlnleuGluThrHisAsnArglyMetValaThr 900
Db 2557 AATAAGAGAGCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2616
```

QY	901	PhelysPheaspLeuaspGlyaspAsnProgluInlealathrIleMetValAspAsn	920	QY	1261	IleGlnLeuSerSerThrSerThrProThrLeuAlaGluThrValValSerAla	1280
Db	2617	TTCAATTTGACCTAGATGATGACAAACCCGAGGAGATAGCAACATTAATGTTAGACAT	2676	Db	3697	ATTGACCTTAGCAGCAGTACTTCACTCACTTACTGTAACCGGTGATGAGCGCA	3756
QY	921	AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu	940	QY	1281	HisSerLeuAspIysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla	1300
Db	2677	GACTTATCTTAGCAATAGAGAGAGAGCTGTTTGATCTCAAGTCGAGAAATTAATGAA	2736	Db	3757	CACCTACTAATAGACATCTCATAGCAGTAACTGATGGCTTCTCCCTCTGCA	3816
QY	941	LysAlaAspGluMetLeuSerGluAspValSerValGluProgluInIleAspGluGlyLeu	960	QY	1301	ProSerSerSerSerSerProgluAlaGlyValSerSerThrIleSerGlnProgluGly	1320
Db	2737	AAAGTCATGAAATGCTAGTAGAGATGTCAGTGGAAACCAAGGGGATCAGGAGATTG	2796	Db	3817	CCATCTTCTCTTCTCTCTCGAGCAGAGTGTCTAGTATATTTCTACGCTGTGTGG	3876
QY	961	GluSerLeuGlnGlyIleAspAspTyrGlyPheSerGlySerGlnIleLeuGluGlyGlu	980	QY	1321	LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla	1340
Db	2797	GAGAGTCACAGAGAAAGATGACTATGGCTTTTCAGGTTCTCGAAATTTGGAAAGAGAG	2856	Db	3877	CTGCATCTTGGTCATTCATCATGATAGCTTCACTCTATTTCTTCCCAAGCAGCA	3936
QY	981	PhelysglnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer	1000	QY	1341	GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro	1360
Db	2857	TTCAACAACCAATCTCGCTCTTCCATGCCACAGCAAAATAGGCATCTTACCAAGTTCT	2916	Db	3937	GGACCTACTTCTACCTTATTAACCCCAAGTACCTGTATCCCACTTGTGTACAGCT	3996
QY	1001	LeuThrGlnValIleHisSerAlaGlyArgArgPheIleValSerProValProGluSer	1020	QY	1361	ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu	1380
Db	2917	TTAAGTCAGATGTTCATTTGCGGAAAGCGGTTTATAGAGTCCGTGCCAGAAAGC	2976	Db	3997	GTTGCCAATGTGCTGCTGTACAGACAGACATTAATTCATAGTCAGCTCAACCGACTTG	4056
QY	1021	ArgLeuArgGluSerIleValPheProSerGluIleThrAspThrValAlaAlaSerThr	1040	QY	1381	LeuProAsnGlnProHisThrHisCysProGluValAspSerAspThrGlnProLysAla	1400
Db	2977	CGATTAGCAGATCAAAAGTTTCCCGCAGTGAATTAACAGATACAGTTGCGCTCTACA	3036	Db	4057	CTTCCCAACCGAGCCCATCTACTATTGCTCTGAATGATTTGTGATACCAACCCAAAGCT	4116
QY	1041	AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnAla	1060	QY	1401	ProGlyIleLeuAspIleIleYsThrLeuGlnGluIleLeuArgSerLeuPheSerGlnHis	1420
Db	3037	GCTCAGAGCCCTGGATGATACCTTGTCTCATCTGCATCATCTTACCTTCAACAGAGCC	3096	Db	4117	CCTGGAATGTGACATTAAGACCTTGAAGAAAGTGGCGTCTGTTCAAGGAACAC	4176
QY	1061	PheSerGluLeuArgArgAlaGluMetThrGluGlyProAsnThrAlaProProAsnPhe	1080	QY	1421	SerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrSerLeuValIleGluSer	1440
Db	3097	TTTCTGAACTTAGCGTGCCCAATGACAGAGAGCAACCAATACAGCACTTCCAAACTT	3156	Db	4177	AGCTCATGTGAGCTCAGCATGCTGTCTCATCTGAGAGACTCACTAGTACATAGAGGC	4236
QY	1081	SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal	1100	QY	1441	ThrValThrProGlyIleProThrThrAlaValAlaProSerIleLeuLeuThrSerThr	1460
Db	3157	AGTCATACAGACCAACATTTCCAGTACTCTCTCTTCTTAAGATGATGCTGAGATC	3216	Db	4237	ACTGTCAACAGGACATCCCACTACTAGCTGTGTCACCAAGCAAACTCTGACTTAC	4296
QY	1101	ProThrThrAlaAlaIleThrAlaProValProAlaThrSerSerProProAsnAspIle	1120	QY	1461	ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr	1480
Db	3217	CCAACCAAGCAGCAGCCACAGCACCACTCCCTGCAACAGAGCCCTCTTAATACACAT	3276	Db	4297	ACAAGTACTGTCTTACCAACCAACCAATTAACCTAGAGAACAGTGTGCGCAGTTACA	4356
QY	1121	SerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGlyVal	1140	QY	1481	ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValIlys	1500
Db	3277	TCCACATAGTAATTCAGTCTGAGGTTCAGGTGCCACTGAAGAGGGATTGCTGAGATT	3336	Db	4357	CCAAGTGTACACCTGGGCAAGTTTCTACCCCACTGACGACTACTACATCAGGAGTAAA	4416
QY	1141	AlaThrSerThrGlyValValThrSerGlyLeuProIleProProValSerGluSer	1160	QY	1501	ProGlyThrAlaProSerIysProProLeuThrIleLysAlaProValLeuProValGlyThr	1520
Db	3337	GCCACCAAGCAGGTGTGTAATCTTCAAGTGTCTCCCATACCACTGTGTCTGAAACA	3396	Db	4417	CCTGGAAGTGTCTCCCTCAAGCCACCTTCACTAAGGCTCGGTGCGCAGTGGTACT	4476
QY	1161	ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr	1180	QY	1521	GluLeuProAlaGlyThrLeuProSerGluGlnLeuProProPheProGlyProSerLeu	1540
Db	3397	CCAGACTTCTCCAGGTGTGTTCAAGTATCCAAATACCTGCAGTGTCTCAATATCTACT	3456	Db	4477	GAACCTTCAGAGGTACTCTTACCAAGCAGAGGTGCCACCTTTTCCAGGACCTTCTCTA	4536
QY	1181	ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla	1200	QY	1541	ThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeuArgArgThrLeuSerPro	1560
Db	3457	ACATCCCGCTCACTCAAGTCCCAACATCCATCTGAGATCGTGTCTTCAAGTACACA	3516	Db	4537	ACCCAGTCCCAAGCACTCTAGAGATCTTATGTCTCAATGAGAAGAACCTTAGTCCA	4596
QY	1201	LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlyIleSerThrAlaThr	1220	QY	1561	Glu**IleThrValThrSerAlaValAlaGlyProValSerMetAlaAlaProThrAlaIle	1580
Db	3517	CTGATACCTTCAAGTAAAGTTTCAAGCACTTCAAGCTTGCAGGGGGAGTAACTGCTACC	3576	Db	4597	GAGATATACAGAGACTTCTGCGGTGTGCTGTCTGTCTCAAGGGGCTCCAAACGACATTC	4656
QY	1221	ProGlyProIysProProAlaValValSerGlnGlnAlaIleGlySerThrThrValGly	1240	QY	1581	ThrGluAlaGlyThrGlnProGlnGlyValSerGlnValIlysGluGlyProValLeu	1600
Db	3577	CCAGGTCTTAAGCCTCCAGCTGTAGTATCTCAGCGCAGCAGCAGCAGCACTACTGTGGA	3636	Db	4657	ACAGAGAGGAAACAGCTTCAAGAGGGTCTTCTCAAGTCAAGAAAGGCCCTGTCTCTA	4716
QY	1241	AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGluLeuSer	1260	QY	1601	AlaThrSerSerGlyAlaGlyValPheIleMetGlyArgPheGlnValSerValAlaAla	1620
Db	3637	GCCACATTAACTCAGTTTCTTACACCACTTCACTTCCAGAGCAGGCTTCCAGCTGTCC	3696	Db	4717	GCAACTAGTTCAGAGCTGGTGTTTTAAGATGGAGATTTACAGGTTTCTGTTGACAGA	4776
QY	3637	GCCACATTAACTCAGTTTCTTACACCACTTCACTTCCAGAGCAGGCTTCCAGCTGTCC	3696	QY	1621	AspGlyAlaGlnIysGluGlyIleLysAsnIysSerGluAspAlaIysSerValHisPheGlu	1640

Db 4777 GACGGTGCACAGAAAGAGGTTAAATAAGTCAGAAAGATGCAAAATCTGTTTCATTTTGA 4836
Qy 1641 SerSerThrSerGlnSerSerValLeuSerSerSerProGlnSerThrLeuVallys 1660
Db 4837 TCCACACCTCAGAGTCTCTCAGTCTATCAAGTAGTCCAGAGATCACTTGCTGAGAA 4896
Qy 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerLahs 1680
Db 4897 CCAGAGCCGAATGGATTAACCATCCCTGGTATCTTTCAGATGTCCAGAGATGCCAAC 4956
Qy 1661 LysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnProThrLysValGlyArgPhe 1700
Db 4957 AAAATACCTGCTCAGAGCAAAAGTCAGACACTGGGAGCCCTACCAAGGTTGACATTTT 5016
Qy 1701 GlnValIleThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnLys 1720
Db 5017 CAGGTGACACACTACAGCAAAAGAGTGGTGGTCTCTGTATCAAAAACAGAGAGCAAG 5076
Qy 1721 IleThrAspThrLysLysGlnGlyProValAlaSerProProPheMetAspLeuGln 1740
Db 5077 ATCACTGACACAAAGAAAGAGACAGTGGCATCTCTCTTATGATTTGAGAAACA 5136
Qy 1741 AlaValLeuProAlaValIleProLysLysGlnLysProGlnLeuSerGlnProSerHis 1760
Db 5137 GCTGTTCTTCTGCTGTATCAAAAGAGAAAGAGAGCTGAACCTGTCAGAGCTTCACAT 5196
Qy 1761 LeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db 5197 CTAAATGGGCGCTCTTGAGCCGAGGCGCTTTTAAAGTAGGAGATGGATATATGT 5256
Qy 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
Db 5257 TCCGGTATCCACACTGCCCCCATCAGCTGAGCTAAAGAGCTTCCTAGCCAGAAATCTA 5316
Qy 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSerSerAspAsnGlnSerAsp 1820
Db 5317 AGTCAAAGCTTAGTAATTCATTTTAATCTCTTCATGAGTACGCAATGAGTCAGAT 5376
Qy 1821 IleGlnAspGlnAspLeuLysLeuGlnLeuArgLysLeuArgLysLeuLysLeuLys 1840
Db 5377 ATCCAAAGTGAAGACTTAAAGTGAAGCTGCGACACATCAAGATAAACATCTCAAAAG 5436
Qy 1841 IleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLeuLysThrLysLeuGly 1860
Db 5437 ATTACAGAGCTCGCAGAGTCCGACAGACATGAATGATCTTTGTATACCAACTGGGC 5496
Qy 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArg 1880
Db 5497 AAGGTGCCCCCTGCTGTAATTAATCCCGCAGCTGCTCCCTTTCAGGAGAAAGACGACA 5556
Qy 1881 ProThrLysSerLysGlnSerLysSerSerArgSerSerSerLeuGlyAsnLysSerPro 1900
Db 5557 CCCACTTAAAGCAAAAGGAGCAAAATCTAGTCAGACACTCTCTTGAGGAAATAAGCCCC 5616
Qy 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
Db 5617 CAGCTTTAGGTAACTGTCTGTGTCAGAGTCCAGTTCAGTTCGACACCCACAGAGCC 5676
Qy 1921 LeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGlnLeuGlnProLeuLys 1940
Db 5677 CTCACACCTCTGCGAACATCCAGAGTCCGCGCAGATCAGCTGTTCACACCCCTTAAAG 5736
Qy 1941 ProSerProSerSerAspAsnLeuLysSerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db 5737 CCATCTCCCTCAGTGAACCTCTATTTCAGCTTTCACAGTGAAGGAGCCATTTCAGTA 5796
Qy 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db 5797 CCAAGCCCTTCTGCTCCAGGTCAGAGAACGACGACGACAAACACTGTTGGGGAACAAGTG 5856
Qy 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000

Db 5857 AACAGCCAGCGCCCAAGCTCAGCTCCTGCACTGATGATGTCAGACAGAGGACATTC 5916
Qy 2001 ThrAspAspLeuHisLysLeuValAspAsnThrAlaArgAspAlaMetAsnLeuSerGly 2020
Db 5917 ACAGATGACTTGACCAAGTGGTAGACAAATGGAGCCCAAGATGACATGATCTTCAGGC 5976
Qy 2021 ArgArgGlySerLysGlyHisMetAsnThrGlnGlyProGlyMetAlaArgLysPheSer 2040
Db 5977 AGAGAGGAGAGCAAAAGGACATGAATTAAGAGGACCTGGAAGAGCAAGAGATTTCT 6036
Qy 2041 AlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGlyLysSerAlaProIleSer 2060
Db 6037 GCACCTGGGCAACTGTGATCATCTCACTGAACTCGAGGAGCTCTGCCCACTCTCT 6096
Qy 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnTyr 2080
Db 6097 GCAGATCAGTACTCTCTGAGTCACTTCAACAGATCATGATGCTGCCCCCAGACAGAT 6156
Qy 2081 GlyPheProAlaThrProPheGlyAlaGlnThrSerSerGlyThrGlyProAlaProGln 2100
Db 6157 GGCCTTCCAGCTACCCCATTTGGCCCTCAATGAGTGGAGCGGTGGCCACACACAG 6216
Qy 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
Db 6217 CCACCTGGCCAGTTTCCAACCTGTGGAGACTGCTCTCTTGCAGAAATTCACATCAGCAAT 6276
Qy 2121 LeuGlnLysSerLysSerAsnProProGlySerAsnLeuArgThrThr 2136
Db 6277 TTGCAAGAAATCCATCAGCAACCCCCAGGCTTCAACCTGCGAGCACT 6324

RESULT 7
ADC99101
ID ADC99101 standard; cDNA: 7152 BP.
XX
AC ADC99101;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human KRP cDNA - SEQ ID 54.
XX
KW anti-HIV; antiallergic; antiinflammatory; antianemic; antiparkinsonian;
KW nootropic; anticonvulsant; antiarteriosclerotic; antiasthmatic;
KW immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;
KW antidiabetic; nephrotropic; antigout; thymimetic; neuroprotective;
KW osteopathic; antiarthritic; antiparasitic; antihelminthic; antispasmodic;
KW uterachic; ophthalmologic; antirheumatic; haemostatic; KRP;
KW cancer; developmental disorder; atherosclerosis; cirrhosis; hepatitis;
KW Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;
KW diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;
KW helminthic infection; transgenic; gene therapy; human; ss; gene.
XX
OS Homo sapiens.
XX
PN W0200303680-A2.
XX
PD 24-APR-2003.
XX
PF 17-OCT-2002; 2002WO-US033723.
XX
PR 19-OCT-2001; 2001US-0345474P.
XX
PR 02-NOV-2001; 2001US-0343910P.
XX
PR 13-NOV-2001; 2001US-0333098P.
XX
PR 16-NOV-2001; 2001US-0332424P.
XX
PR 30-NOV-2001; 2001US-0334288P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Bandman O, Baughn MR, Becha SD, Borowsky ML, Duggan BM;
PI Emerling BM, Foreyche IU, Gandhi AR, Gorvad AE, Griffin JA;
PI Gururajan R, Hafalia AJA, Khan FA, Ial PG, Lee SY, Lee SY;
PI Lindquist EA, Lu DAM, Lu Y, Marquis JP, Nguyen DB, Arvizu CS;

PI Ramkumar J, Recipon SA, Richardson TW, Swarnakar A, Tang YT,
PI Thornton MB, Tran UK, Chawla NK, Warren BA, Yang J, Yao MG, Yue H;
PI Zebajadian Y;
XX
XX WPI: 2003-403214/38.
DR P-PSDB; ADC99049.

XX New human kinases and phosphatases and polynucleotides, useful for
PT diagnosing, treating or preventing autoimmune or inflammatory disorders
PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
PT cancer or hepatitis.

XX Claim 5; SEQ ID NO 54; 424pp; English.

XX The invention relates to a novel isolated polypeptide which is a human
CC kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing cell proliferative disorders such as atherosclerosis,
CC cirrhosis, hepatitis and cancer, developmental disorders e.g. mental
CC retardation, neurological disorders including Alzheimer's disease and
CC Parkinson's disease, autoimmune and inflammatory disorders such as
CC Crohn's disease and diabetes mellitus and finally, viral, bacterial,
CC fungal, parasitic, protozoan or helminthic infections. Furthermore, the
CC polynucleotides encoding KPP may be useful for creating transgenic
CC animals to model human disease, as well as during gene therapy
CC procedures. The current sequence is that of the human KPP cDNA of the
CC invention.

XX Sequence 7152 BP; 1915 A; 1952 C; 1682 G; 1603 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	7152
Score: 0	Matches:	2107
Percent Similarity: 10646.00	Conservative:	0
Best Local Similarity: 98.64%	Mismatches:	1
Query Match: 98.46%	Indels:	28
DB:	Gaps:	1

US-10-010-720-14 (1-2136) X ADC99101 (1-7152)

QY 1 MetSerGlyGlyAlaAlaGlnLysGlnSerThrProGlySerLeuPheLeuSerPro 20
DB 361 ATGCTGGCGGCGCGGAGAGACGAGACACTCCCGGTCCCTGTTCTCTCGCGG 420
QY 21 ProAlaProAlaProLysAsnGlySerSerSerSerSerValGlyGlnLysLeuGly 40
DB 421 CCGGCTCTGCGCCCGAGAAATGGCTCCAGCTCCGATCTCTCGTGGGGAAGAACTGGA 480
QY 41 AlaAlaAlaAlaAspAlaValThrGlyArgThrGlnGlyArgArgArgHisThr 60
DB 481 GCCGGCGCGCGGAGCTGTGACCGCGGAGACCGAGAGTAAAGCGCGCGCACACT 540
QY 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrThrGlnHisArgPhePhe 80
DB 541 ATGGAACAAGACAGCCGCGGGCGCGCGGACCACTACACACAGACAGACCGCTTCTTC 600
QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100
DB 601 CGCGGAGCGTCAATTCGCACTCCATGCACTGAGACTTCCGAGCTTCTCTT 660
QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGlnProHis 120
DB 661 TCCCTGCGCCAGCCAGCATCCCGCGGCTGTCCCGGAGAGTGTCCACCGGAGCCAC 720
QY 121 ArgGlnGlnThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
DB 721 CGGAGAGAGACCGGAGACCGCACCGCCACTTCCAGGTAGCCACAGACCTCCAGCCCT 780
QY 141 AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
DB 781 GCCGCGCTCGGGGAACAGGCGCGCGGCGCTGCGCCCTCGACTGTCCACAGCATGCC 840
QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGlnProPro 180

DB 841 AGCAAAACCGCCAGAGTCCAGCCTAGGCTTGTGGGAGCAAAAGAGAGCCCGCGG 900
QY 181 AlaArgSerGlySerGlyGlySerAlaLeuGlnProGlnGlnGlnLysSerGlnGln 200
DB 901 GCGAGAGTGCAGCGCGCGCGGAGCGCCAGGAGCAACAGAGAAAGGAGCCAGCAG 960
QY 201 GlnAspAspIleGlnGlnLeuGlnThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
DB 961 CAGGATATATCGAAGAGCTGAGAGCCAGGCGGTGGGAAATGTAAAGATGCGCGCTT 1020
QY 221 LeuLysPheAspIleGlnIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
DB 1021 CTCAGATTGACATCGAAATCGCAGAGCTCTTTAAGACGGCTCAAAAGGCTCTGAC 1080
QY 241 ThrGlnThrThrValGlnValAlaTyrCysGlnLeuGlnAspArgLysLeuThrLysSer 260
DB 1081 ACTGAAACCAACCGGAGAGTGCCTGCTGTGAATCGAGATCGAAATTTAAACAAAGCT 1140
QY 261 GlnArgGlnArgPheLysGlnGlnAlaGlnMetLeuLysGlyLeuGlnHisProAsnIle 280
DB 1141 GAGAGGAGAGATTTAAAGAAAGCTGAATGTTAAAGGCTTCAGCATCCCAATATT 1200
QY 281 ValArgPheTyrAspSerTyrGlnSerThrValLysGlyLysLysCysIleValLeuVal 300
DB 1201 GTTAGATTATATGATTCCTGGGAATCCACAGTAAAGAAAGAGTGCATGTTGTGTG 1260
QY 301 ThrGlnLeuMetThrSerGlyThrLeuLysThrTyrLeuLysAspPheLysValMetLys 320
DB 1261 ACTGAACCTTATGAGCTGTGAACCTTAAACGATATCGAAAGGTTTAAAGTATGAG 1320
QY 321 IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
DB 1321 ATCAAAATTTCAAAAGCTGGTCCGTCAGATCTTAAAGGCTCTTCACTTCAATCT 1380
QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
DB 1381 CGAACTCCACCTATCATTCACCGGATCTTAAATGTACAAACATCTTATACCGGCGCT 1440
QY 361 ThrGlySerValLysIleGlyAspLeuGlnLeuAlaThrLeuLysArgAlaSerPheAla 380
DB 1441 ACTGGCTCGTCAAGATTTGAGACTCGTGTGGCAACCTGAAAGCGGCTTCTTTTGC 1500
QY 381 LysSerValIleGlyThrProGlnPheMetAlaProGlnMetTyrGlnGlnLysTyrAsp 400
DB 1501 AAGAGTGTGATAGTACCCAGAGTTTCATGGCCCTGAGATGTAGAGAAATATGAT 1560
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlyTyr 420
DB 1561 GAATCCGTTGACGTTTATGCTTTTGGGATGTGATGCTTGAGATGGCTACATCGAAAT 1620
QY 421 ProTyrSerGlnCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
DB 1621 CTTTACTCGAGTCCAAAATGTGCGCAAGTTCACGTCGCGGACAGTGGGAGTGAAG 1680
QY 441 ProLasePheAspLysValAlaIleProGlnValLysGlnIleGlnGlyCysIle 460
DB 1681 CCAAGCGATTTTGAACAAGTACATTCCTGAAATGTAAGAAATTAATTGAAGATGATA 1740
QY 461 ArgGlnAsnLysAspGlnArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
DB 1741 CGACAAAACAAGATGAAGATTTTCATCAAAACCTTTTGAACATGCTTCTTCCAA 1800
QY 481 GlnGlnThrGlyValArgValGlnLeuAlaGlnGlnLysAspGlyGlnLysIleAlaIle 500
DB 1801 GAGGAACAGAGTACGGGTGAAATTTAGCAAGAAAGATGATGAGAAATTAATGCCATA 1860
QY 501 LysLeuThrPheArgIleGlnAspIleLysLysLeuLysGlyTyrTyrLysAspAsnGln 520
DB 1861 AATTTATGCGTACGATATTGAAGATATTAAAGAAATTAAGGAAATACAAAGATATGAA 1920
QY 521 AlaIleGlnPheSerPheAspLeuGlnArgAspValProGlnAspValAlaGlnGlnMet 540

Dh 1921 GCTATTGAGTTTCTTTGATTAGAGAGATGTCCTCCAGAGATGTTGACACAGAAATG 1980
QY ValGlnSerGlyThrValCysGlnGluIysAspHisIleThrMetAlaIleValAlaIleIysAsp 560
Dh 1991 GTAGAGTCTGGGTATGCTGTGTAAGAGTATGATCAAGACCATGGCTTAAGCTATCAAAAGC 2040
QY ArgValSerLeuIleIysArgIysArgGlnArgGlnArgGlnLeuValArgGlnGlnGln 580
Dh 2041 AGAGTATCATTTATTAAGAGAAACAGAGACGGCGAGTTGTTGTTGCGGAGAGAGAAAGAA 2100
QY LysIleIysGlnGlnGlnSerSerLeuIysGlnGlnValGlnGlnSerSerAlaSerGln 600
Dh 2101 AAAAAAGCAGGAGAGAGAGCTCTCAACAGCAGGTAGAAACATCCAGTCTCCAG 2160
QY ThrGlyIleIysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Dh 2161 ACAGGAATCAAGACCTCTCTCTCTGACACGGGACATCACTCACTTCTTCAACCTTCA 2220
QY AlaSerValSerThrGlnValGlnProGlnGlnProGlnAlaAspGlnHisGlnGlnLeu 640
Dh 2221 GCTTCAGTTTCTACACAGTAGAACTGAAAGACCTGAGGACGATCAACATCAACACTA 2280
QY GlnThrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Dh 2281 CAGTACCAAGCAACCGATATCTGTGTATCTGATGGAGCGTTGACAGTGTCAAGGGA 2340
QY SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerThrGlySerGln 680
Dh 2341 TCCTCTGCTTCAACAGATCTGAGTGAAGCAGCCACAGACAGATTCATATGAGTCCCA 2400
QY HisGlnGlnAlaHisSerThrGlyThrValProGlnHisIleProSerThrValGlnAla 700
Dh 2401 CATGAACAGGACATTTACAGGACAGTCCAGGCGATATACCTTACTGCTCCAAACA 2460
QY GlnSerGlnProHisGlyValIleThrProProSerSerValAlaGlnGlyGlnSerGln 720
Dh 2461 CAGTCTACGCCCGCATGGGGTATCCACCTCCAAAGTGTG----- 2499
QY GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Dh 2499 ----- 2499
QY GlnGlnGlnGlnIleGlnGlnThrAlaProProGlnGlnThrValGlnIleSerLeuSer 760
Dh 2500 ---CAGCAGGGAATACAGACAGACAGCCCTCCTCAACAGACAGTATTAACCTTCA 2556
QY GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Dh 2557 CAGACATCAACCTCCAGTGAAGGCCACTGACACGCGCAGTGAAGTCAACCTCAAGCTCCA 2616
QY GlnValLeuProGlnValSerAlaGlyIysGlnSerThrGlnGlnValSerGlnValAla 800
Dh 2617 CAAGTCTTGCTCAAGTATCAAGCTGGAAGAACAGAGTACTCAGGAGAGTCTCCAGAGTGTCT 2676
QY ProAlaGlnProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Dh 2677 CCTGCAAGCGCAGTTGACAGTGAAGCCAGCCCAAGTCAACCCAGCCAGCACTTGGCTTCC 2736
QY SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGlnAsnVal 840
Dh 2737 TCTGTAGACATGACATTCAGATGTGCTTCAGGTATGAGTGAATGGCAATGAAGAACGTC 2796
QY ProSerSerSerGlyArgHisGlnGlyIleArgThrThrIleIysArgHisGlyArgIleValSer 860
Dh 2797 CCATCTTCCATGGAAGGCATGAAGGAAGAACTCAAAACGGCATTTACCGAAAAATCTGTA 2856
QY ArgSerArgSerArgHisGlnIleThrSerArgProIysLeuArgIleLeuAsnValSer 880
Dh 2857 AGGAGTGGCTCTCGACATGAAAAAACTTCAAGCCCAAAATTAAGAAATTTGAATGTTCA 2916
QY AsnIleIysIysPargValValGlnCysGlnLeuGlnIleThrHisAsnAlaGlyMetValThr 900
Dh 2917 AATTAAGGAGACCGAGTAGTAGAATGTCAATTAGAGACTCATTAATAGAAAAATGGTTACA 2976

QY PheIysPheAspLeuAspGlyIysAspAsnProGlnGlnIleAlaThrIleMetValIleAsn 920
Dh 2977 TTCAAAATTTGACCTAGATGAGTGAACAACCCGAGGAGATGCAACAAATATATGTTGAACAAT 3036
QY AspPheIleLeuAlaIleGlnArgGlnSerPheValAspGlnValArgGlnIleIleGln 940
Dh 3037 GACTTATTTACGAAATGAGAGAGAGTCTTTTGTGATCAAGTCCAGAAATTTATTTGA 3096
QY LysAlaAspGlnMetLeuSerGlnAspValSerValGlnProGlnGlyAspGlnGlnLeu 960
Dh 3097 AAAGCTGATGAATAGCTCAGTAGAGATGTCAGTGGAACCAAGAGGATGATCAGAGGATTG 3156
QY GlnSerLeuGlnGlnIleIysAspAspThrGlyPheSerGlySerGlnIleLeuGlnGln 980
Dh 3157 GAGAGTCTACAGAGAAAGAGTACCTATGCTTTTCAAGGTTCTCAAGAAATGGAGAGAG 3216
QY PheIysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
Dh 3217 TTCAACGACCAATTCCTGCGCTTCCATGCGCAACGCAATAGGCATTCCTACCACTTCT 3276
QY LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGlnSer 1020
Dh 3277 TTAACCTCAAGTTGTTCAATTCTGCGGAAAGCGGTTTATAGAGATCCTGCTCCAGAAAGC 3336
QY ArgLeuArgGlnSerIysValPheProSerGlnIleThrAspThrValAlaAlaSerThr 1040
Dh 3337 CGATTACAGAGATCAAAAGTATTCCTCCAGTGAATTAACAGTATACAGTGTGCTGCTTACA 3396
QY AlaGlnSerProGlnMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAla 1060
Dh 3397 GCTCAGAGCCCTGGAATGAATTCCTGCTCAGCTGCAATCATCTTACTGATCAACAGGCC 3456
QY PheSerGlnLeuArgArgAlaGlnMetThrGlnGlyProAsnThrAlaProProAsnPhe 1080
Dh 3457 TTTTCTCAACTTACAGCTGCCCAATGACAGAAAGACCAATACAGCACTCCAAACTTT 3516
QY SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGlyVal 1100
Dh 3517 AGTCATACAGGACCAACATTCACAGTGAAGTACCTCTTTCTTAATATACATTTGCTGAGTCT 3576
QY ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
Dh 3577 CCAACCAAGAGAGACGACACAGCACTCCCTGCAACAGCAAGCCCTCTCAATAGCAATT 3636
QY SerThrSerValIleGlnSerGlnValThrValProThrGlnGlnGlyIleAlaGlyVal 1140
Dh 3637 TCCACATCAGTAATTCAGTCTGAGGTACAGTGCCTGAAAGAGGGGATGCTGAGATT 3696
QY AlaThrSerThrGlyValValThrSerGlyIleuProIleProProValSerGlnSer 1160
Dh 3697 GCCACGACACAGAGTGTGTAATTCAGGTGTGTCCCATACCACTGATGTGTAANCA 3756
QY ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
Dh 3757 CCAAGTACTTTCCAGCGTAGTTTCAAGTATCAACATTAACCTGCAAGTGTCTCAATTAAT 3816
QY ThrSerProSerLeuGlnValProThrSerThrSerGlnIleValValSerSerThrAla 1200
Dh 3817 ACATCCCGGACACTTCAAGTCCCAACATCAACATCAAGATCGTTGTTTCAAGTACAGA 3876
QY LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlyIleThrAlaThr 1220
Dh 3877 CTGATCTCTTCAAGTACAGTTTCAAGCACTTCACTCTGACGGGAGCAGTACTGCTACC 3936
QY ProGlyProIysProProAlaValValSerGlnGlnAlaAlaGlySerThrThrValGly 1240
Dh 3937 CCAAGTCTTAAGCTCCAGCGTAGTATCTCAGAGCAAGCAAGCACTACCTGTGGA 3996
QY AlaThrIleThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Dh 3997 GCCACATTAACATCAGTTTCAACACCACTTATTCACAGCAAGCTTCAAGCTGTCC 4056

Db	901	ACTGACTTATGACGCTCGGAACACTTAAAAAGCATCTGAAAAAGTTTAAAGATGAAAG	960
Qy	321	ILEYVALLEUAARGSERIPCYAARGINILEULYSGLYLEUNGINPHELEUHSIR	340
Db	961	ATCAAGATTCTAAGAGAGCTGGTGGCGTCAGATCTCTTAAAGGTCTTCAGTTTCTTCATCT	1020
Qy	341	ARGTHRPROFOLLEILEHSARGASPLEULYSYASPSANILEPHELLETHRGLYPRO	360
Db	1021	CGAACTCCACTATCATCTTCAACCGGGAATCTTAAAGTACAAACATCTTTATCACCGGCGCTT	1080
Qy	361	THRGYSERVALVLSILEGLYASPLEUGLYLEULATHRLEULYSARGALASERPHENLA	380
Db	1081	ACTGGCTCAGCAAGATTGAGACCTCGGTGTGGCAACCTTGAGCGGCGCTTCTTTGGC	1140
Qy	381	LYSSERVALILEGLYTHRPROGLUPHEMELAPROGLUMETTYRGUGULYSYTYASP	400
Db	1141	AAGAGTGGATAGGTATCCCGCAGATTCATGGCCCCCTGAGATGTATGAGGAATATGAT	1200
Qy	401	GLUSERVALASPVALLYRALPHEGLYMETCYSMETLEUNGULELATHRSERGLYTR	420
Db	1201	GAATCCGTTGACGTTTATGCTTTTGGGATGTGCATGCTTGAGATGGCTCATCTCGAATAT	1260
Qy	421	PROTYRSERGLUCYSGINASNALALGLNILEYTHRRATRYVALTHRSERGLYVALYS	440
Db	1261	CCTTACTCGAGGTGCCAAATAGCTGCACAGATCTACCGTGCCTGACCACTGGGGTGAAG	1320
Qy	441	PROLASERPHEASPLYVALALALEPROGLUVALLYSGULLEGLUGLYCYAILE	460
Db	1321	CCAGCCAGTTTGGACAAAGTGACAAATCTCTGAACTGAAGGAATATTATGGAAGATGATA	1380
Qy	461	ARGGINASULYASPGULARGTYRSERILEYASPLEULEUASNIISALAPHEPHEGLN	480
Db	1381	CGACAAAAACAAAGATGAAGAAAGATATTCCTACAAAGACCTTTGAACCATGCTTCTCCAA	1440
Qy	481	GLUGLUTHTRGLYVALARGVALGULLEULALGLULUSPASPGLYGLULYSILEALALE	500
Db	1441	GAGGAACACAGGATACGGGTAGAAATTACCAAGAAAGATGATGAGAAAAAATAGCCATA	1500
Qy	501	LYSLEUTRPLEUARGILEGLUASPIILEYSLYSLEULYSGLYTYLYSASPASNGLU	520
Db	1501	AAATTATGGCTACGATTTGAAGATTTTAAAGAAATTTAAAGGAAAAATACAAAGATATGAA	1560
Qy	521	ALAILEGLUPHESERPHEASPLEUNGULARGASPVALPROGLUASPVALALEGLINUMET	540
Db	1561	GCTATTGAGTTTCTTTTGATTAGAGAGAGATGTCCAGAAAGTGTGGCACAGAATG	1620
Qy	541	VALIGUSERGLYTYRYVALCYSGULGLYSPHSILYSTHRECALALYSALALEYASP	560
Db	1621	GTAGAGTCTGGGTATGCTGTGMAAGSTATACAAAGACATAGGCTTAAAGCTATAAAGAC	1680
Qy	561	ARGVALSERLEULEYASRGULYSARGULNARGINLEULVALARGGLUGLUGLNU	580
Db	1681	AAGATATCTTAATTAAAGGAAACGAGACGCCGCGCTTGGTACGGGAGGACCAAGAA	1740
Qy	581	LYSLYSGLINGLUGLUSERSERLEULYSGLINGLUNALGLINGUSERSERALASERGLN	600
Db	1741	AAAAAAAACGAGAAAGAGAGCAGCTCTCAACACGACGATGAACAATCCAGGTCTCCAG	1800
Qy	601	THRGYILLEYGINLEUPROSERALASERTHRGLYLEPROTHRALASERTHTRHSER	620
Db	1801	ACAGGATATMACAGCTCCCTTCTCTCTGACACCGGCATATCTTACGCTTCTTCAACTTCA	1860
Qy	621	ALASERVALSERTHGLINVALGIUPROGLUGLUPROGLUVALASPGLNHISGLINLEU	640
Db	1861	GCTTCAGTTTCTACACAGATGAACTTGAAAGACTTGAGCGCATCAACATCAACACTA	1920
Qy	641	GINTYRGINGINPROSERTLESERVALLEUSERASPGLYTHRVALASPSERGLYGLNLY	660
Db	1921	CAGTACCAACACCCAGTATATCTGTGTTATCTGATGGGACGGTGTGAACAGTGTACAGGA	1980
Qy	661	SESRSERVALPHERTHGLUSERARGVALSERSERGININTHRVALSERTYGLYSERGLN	680
Db	1981	TCTCTGTCTTACAGAAATCTGAGTGAACACCAACAGACAGTTTCAATATGTGTCCCA	2040

QY	681	HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla	700
Db	2041	CATAAACAGGACATTCCTACAGGACACAGTCCAGGGCATATACCTCTCTACTGTCACAGCA	2100
QY	701	GlnSerGlnProHisGlyValIlyProProSerSerValAlaGlnGlyGlnSerGlnGly	720
Db	2101	CAGTCTACAGCCCATGGGGTATATCCACCTCAAGTGTG-----	2139
QY	721	GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740
Db	2139	-----	2139
QY	741	GlnGlnGlnGlyIleGlnGlnInThrAlaProProGlnGlnThrValGlnIlySerLeuSer	760
Db	2140	---CAGCAGGGAAATACAGACAGACAGCCCTCTCTCAACAGACAGAGATATCACTTTCA	2196
QY	761	GlnThrSerThrSerSerSerGlnAlaIleThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	2197	CAGACATCAACTCCAGTGAAGGCCCACTACTGACACAGCACTGATGATGAGCTCAAGCTCCA	2256
QY	781	GlnValLeuProGlnValSerAlaGlyLysGln-----	791
Db	2257	CAAGCTTGCTGCCTCAAGATACAGCTGGAAACAGGGCTTCCCACTCGACTGCCAACAG	2316
QY	791	-----	791
Db	2317	TACCCAGAGATTCAAAATATTGCTCCCTCTTCCAAAGTGGCTCTGTTCATCCATCTT	2376
QY	791	-----	791
Db	2377	ACAGTCTAMCCCTCCCATGCCAGACAGAAAGTACTGGCTACACTCGGTGACTTTCCACA	2436
QY	791	-----	791
Db	2437	GTGGTGACGCTTATGTGGAAATCAAAATCTTTAAGTTCCTATGGGTGTGAGAGAGACAG	2496
QY	791	-----	791
Db	2497	GTTCAAGTGTCCAGCCAGACGAGGGAGTTTACACAAAGCCCCCACTACATCCTCCACAA	2556
QY	792	-----SerThrGlnGlyValSerGlnValAlaProAlaGluProValAlaVal	807
Db	2557	GCAGTTTTGAGAGATACACAGGAGGTCTCTCAGTCTCTCTCGACAGCCAGTTGCAGTA	2616
QY	808	AlaGlnProGlnAlaIleThrGlnProThrThrLeuAlaSerSerValAlaSerAlaHisSer	827
Db	2617	GCACAGCCCAAGTACCCACGCCACCACTTTGGCTTCTCTGTGACAGTGACATTTCA	2676
QY	828	AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyAlaGHis	847
Db	2677	GATGTGTGCTTCAGGTATAGATGATGGCANTAGAACGTCATCTTCAGATGGAAGGCAT	2736
QY	848	GluGlyValArgThrThrLysAlaGHisIleValArgLysSerValAlaSerAlaArgSerAlaGHis	867
Db	2737	GAAAGAGAACTACAAACCGCATTTACCGAAATCTGTGACAGATCGCTCCACATGAAG	2796
QY	868	LysThrSerArgProLysLeuArgIleLeuAsnValSerAsnLysGlyAspArgValVal	887
Db	2797	AAACATTCACGCCCAAAATTAAGATTTTGAATGTTTCAAAATTAAGAGGACCGAGTGTGA	2856
QY	888	GluCysGlnLeuGluThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly	907
Db	2857	GATGTCAATTGAGAGCTCATATATGAGAAATGTTTACATTCMAATTTGACTTGATGATGT	2916
QY	908	AspAsnProGluGluIleAlaThrIleMetValAsnAsnAspPheIleLeuAlaIleGlu	927
Db	2917	GACAAACCCGAGAGATAGCAACATTTGTGCAACATGACTTATCTCTGCAAATAGAG	2976
QY	928	ArgGluSerPheValAspGlnValArgIleIleGlnLysAlaAspGluMetLeuSer	947
Db	2977	AAGAGATCGTTTGATGATCAAGTGGCAAAATATATGAAAAAGCTGATGAATGCTCAGT	3036

QY 948 GluAspValSerValGluProGluGlyAspGluGlyLeuGluSerLeuGluGlyLysAsp 967
 Db 3037 GAGGATGTCAGTGTGGAAACAGAGGGTGTATCAGGATTTGGAGCTTACAGAGAAAGAT 3096
 QY 968 AspTyrGlyPheSerGlySerGlnLysLeuGluGlyGluPheLysGlnProIleProAla 987
 Db 3097 GACTATGGCTTTTCAGGTTCTCAGAAATTTGAAAGGAGATTCAAACCAACCAATTCCTGGC 3156
 QY 988 SerSerMetProGluGlnIleGlyIlePheThrSerSerLeuThrGluValValHisSer 1007
 Db 3157 TCTTCATCCACAGCAAAATAGCATTCCTACCAAGTTCTTAACCTCAAGTTTTCATCTT 3216
 QY 1008 AlaGlyArgArgPheIleValSerProValProGluSerArgLeuArgGluSerLysVal 1027
 Db 3217 GCGGGAAGCGGCTTTATAGTGAAGTCTCTGCGCAAAAGCCGATTACGAAATCAAAAGTT 3276
 QY 1028 PheProSerGluIleThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
 Db 3277 TTCCCGAGTGAATACAGANTACAGTTGCTGCTTACAGCTCAGAGCCCTGGAAATGAAC 3336
 QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAla 1067
 Db 3337 TTGTCTACTCTGCATCATCTCTTAGTCTACACAGGCTTTTCTGAACCTTAGAGCTGCC 3396
 QY 1068 GlnMetThrGluGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
 Db 3397 CAATGACAGAAAGGCCCAAAVACAGCACCTCCAAACTTAGTCATACAGAACCAATTT 3456
 QY 1088 ProValValProProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaThr 1107
 Db 3457 CCAAGTAGTCTCTCTTTAAGTAGCATTTGCTGAGTCCCAACCAAGCAGCAGCCCA 3516
 QY 1108 AlaProValProAlaThrSerSerProProAsnAspIleSerThrSerValIleGlnSer 1127
 Db 3517 GCACAGTCCCTGCAGAACAGACGCCCTCTATATGACATTTCCATCTCAGTAATTCAGTCT 3576
 QY 1128 GluValThrValProThrGluGluGlyIleAlaGlyValAlaThrSerThrGlyValVal 1147
 Db 3577 GAGGTACAGTGCCTCAGAGAGGGGATGCTGGAGTTGCCACAGACAGGTGTGGTA 3636
 QY 1148 ThrSerGlyGlyLeuProIleProProValSerGluSerProValLeuSerSerValVal 1167
 Db 3637 ACTTCAGGTGGTCTCCCAATACACCTGTGTCTGATATCACAGTACCTTCCAGGGTGT 3696
 QY 1168 SerSerIleThrIleProAlaValAlaValSerIleSerThrThrSerProSerLeuGlnVal 1187
 Db 3697 TCAAGTATCAAAATCTCGAGTGTGTCTCAATATCTACTATCATCCCGTCACTTCAGATC 3756
 QY 1188 ProThrSerThrSerGluIleValValSerSerThrAlaLeuTyrProSerValThrVal 1207
 Db 3757 CCCACATCCACATCTGAGATGTTGTTCTAGTACAGCATGTATCTTTCAGTAACAGTT 3816
 QY 1208 SerAlaThrSerAlaSerAlaGlyGlySerThrAlaThrProGlyProLysProProAla 1227
 Db 3817 TCAGCAACTTAGGCTCTGCGGGGGGAGTAGTCTGATCCCAAGGTCTTAAACCTCCAGCT 3876
 QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
 Db 3877 GTACTATCTCAGCAGGAGCAGCAGCATCTGTGGAGGACCATTAACATTCAGTTTCT 3936
 QY 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerThr 1267
 Db 3937 ACCACCACTTCATCTCCAGAGCAGCTTTCACAGCTGTCCATTCAGCTTAGAGAGTACT 3996
 QY 1268 SerThrProThrLeuAlaGluThrValValValSerAlaHisSerLeuAspLysThrSer 1287
 Db 3997 TCTACTCTACTTAGCTGAAACCGTGTAGTTAGCCACACATCACTAAGATTAACATCT 4056
 QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
 Db 4057 CATAGCAGTACAACTGATTTGGCTTTCTCCCTCTCTGACACATCTTCTTCTCTCTCT 4116
 QY 1308 GlyAlaGlyValSerSerTyrIleSerGlnProGlyGlyLeuHisProLeuValIlePro 1327
 Db 4117 GGAGCAGAGTGTCTAGTTATATTTCTCAGCCTGTGGCTCCATCTTGTGATCTTCCA 4176
 QY 1328 SerValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
 Db 4177 TCAGTATAGTCTTCTACTCTCTATTTCTTCCCAAGCAGCAGGACCTTCTACACTTTA 4236
 QY 1348 LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal 1367
 Db 4237 TTACCCCAAGTACATAGATATCCACCTGTGTAGACGCTGTGTGCCAATGTGCTGTGTA 4296
 QY 1368 GlnGlnThrLeuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisSer 1387
 Db 4297 CAGCAGACATTAATTCATATGTCAGCTTCAACAGGCTTTGTCTCCACACAGCCCATACT 4356
 QY 1388 HisCysProGluValAspSerAspThrGlnProLysAlaProGlyIleAspAspIleLys 1407
 Db 4357 CATTTGCTGTAAGTGAATTTGATACACACCAAGAGCTCTGGATTTGATGACATTAAG 4416
 QY 1408 ThrLeuGluGlyLysLeuArgSerLeuPheSerGluHisSerSerSerGlyAlaGlnHis 1427
 Db 4417 ACTTAGAAGAAAGCTGCGGTCTCTGTCAGTGAACACAGCTCATCTGAGCTCAGCAT 4476
 QY 1428 AlaSerValSerLeuGluThrSerLeuValIleGlnSerThrValThrProGlyIlePro 1447
 Db 4477 GCCTCTGCTACCTGGAGACCTCACTAGTCAATAGAGACCTGTACACAGGACATCCCA 4536
 QY 1448 ThrThrAlaValAlaProSerSerLysLeuLeuThrSerThrThrCysLeuProPro 1467
 Db 4537 ACTAGTCTGTGGACCAAGCAAACTCCTGACTTACACAAAGTATGTTGTTACCA 4596
 QY 1468 ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValValThrProGlyGln 1487
 Db 4597 ACCAATTTACACATGAGAAACGTTGCTTGCAGTTTACACAGAGTGCACACCTGGGCA 4656
 QY 1488 ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys 1507
 Db 4657 GTTTCACCCAGCAGACATCTACATCAGAGATGAAACCTGGAACTGCTCCTCCAG 4716
 QY 1508 ProProLeuThrLysAlaProValLeuProValGlyThrGluLeuProAlaGlyThrLeu 1527
 Db 4717 CCACCTTAATAAGCTCCGCTCGCTCCAGTGGGTACTGAACCTTCCAGCAGGTACTCTA 4776
 QY 1528 ProSerGluGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnProLeu 1547
 Db 4777 CCAAGCAGAGAGCTGCCACCTTTCCAGAGCTTCTTAAACCAAGCTCCAGCAACTCTTA 4836
 QY 1548 GluAspLeuAspAlaGlnLeuArgArgThrLeuSerProGlu**IleThrValThrSer 1567
 Db 4837 GAGGATCTTATGTCTCAATTTGAGAGAAACCTTAGTCCAGATKATACAGTACTTCT 4896
 QY 1568 AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGluAlaGlyThrGlnPro 1587
 Db 4897 GCGGTGTGCTCTGTCTCAGTCCAGGCTCCAAAGCAATCACAGAAAGCAGAACAGCTT 4956
 QY 1588 GlnLysGlyValSerGlnValLysGluGlyProValLeuAlaThrSerSerGlyAlaGly 1607
 Db 4957 CAGAGGAGTGTCTCAGTCCAAAGAGAGGCTGTCTTACCAACTATTTACAGAGCTGT 5016
 QY 1608 ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyAlaGlnLysGluGly 1627
 Db 5017 GTTTTAAAGATGGAGCATTTTCAGTGTCTGTTCACAGACAGGCTGCCGAAAGAGGCT 5076
 QY 1628 LysAsnLysSerGluAspAlaLysSerValHisPheGlnSerSerThrSerGluSerSer 1647
 Db 5077 AAAAATAGTCAAGATGCAAAAGTGTGTCATTTTCAATCCAGACCTCAGAGTCTCTCA 5136
 QY 1648 ValLeuSerSerSerSerProGluSerThrLeuValLysProGluProAsnGlyIleThr 1667
 Db 5137 GTCTATCAAGTGTGATCCAGAGATCACTTGTGTGAAACAGAGCCGATGGCATTAAC 5196
 QY 1668 IleProGlyLysSerSerAspValProGluSerAlaHisLysThrThrAlaSerGluAla 1687

Db 5197 ATCCCTGGATCTCTTCAGATGTCAGAGAGTGCCCAAACTACTGCTCAGAGCA 5236
 Qy 1688 LysSerAspThrGlyGlnProThrLysValGlyArgPheGlnValThrThrAlaAsn 1707
 Db 5257 AAGTCAGACTGGGCGAGCTTACCAGAGTTGAGCTTTTCAGGGTGAACAATCAGCAAC 5316
 Qy 1708 LysValGlyArgPheSerValSerLysThrGluAspLysIleThrAspThrLysLysGlu 1727
 Db 5317 AAGTGGGTGCTTCTCTGATCAAAAATGAGAGCAAGATCACTGACACAAAGAAAGAA 5376
 Qy 1728 GlyProValAlaSerProProPheMetAspLeuGlnAlaValLeuProAlaValIle 1747
 Db 5377 GAGACAGTGGATCTCTCTTATGATTTGAAACAGCTGTTCTTCCGCTGTATA 5436
 Qy 1748 ProLysLysGlyLysProGlnLeuSerGluProSerHisLeuAsnGlyProSerSerAsp 1767
 Db 5437 CCAAGAGAAAGAGAGCTGAACTGTCAGAGCTTCACTTAATAGGCGCTTCTTGAC 5496
 Qy 1768 ProGluAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHisSerPro 1787
 Db 5497 CCGAGAGCGCTTTTAAAGTAGGAGTGTGATGATGGTTCCGGTAGCCACACTCGGCC 5556
 Qy 1788 HisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer 1807
 Db 5557 CATCAGCTGAGCTCAAGAGCTTCTCAGCCAGATCTAACTCAAGGCTTATGATTTCA 5616
 Qy 1808 PheAsnSerSerLysSerSerSerAspAsnGluSerAspIleGluAspGluAspLeuLys 1827
 Db 5617 TTTACTCTCTTACATGATGATGAGCAATAGTCAGATATCGAAGATGAAAGACTTAAAG 5676
 Qy 1828 LeuGlnLeuArgArgLeuArgAspLysHisLeuLysGluIleGlnAspLeuGlnSerArg 1847
 Db 5677 TTAAGAGCTGGAGCAGCATACGATTAACATCTCAAGAGATTCAGGACTCGACAGTGC 5736
 Qy 1848 GlnLysHisGluIleGluSerLeuThrLysLeuGlyLysValProProAlaValIle 1867
 Db 5737 CAGAAAGCATGAATGATCTTTGATATCCAACTGGGCAAGGTCGCCCTGCTGTTATT 5796
 Qy 1868 IleProProAlaAlaProLeuSerGlyArgArgArgProThrLysSerLysGlySer 1887
 Db 5797 ATTCCTCCAGCTGCTCCCTTCAGGGAGAAAGACACACCACTAAAGCAAAAGGACGC 5856
 Qy 1888 LysSerSerArgSerSerSerLeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSer 1907
 Db 5857 AAATCTAGTCAGAGAGTTCCTTGGGGAATAAAGCCCCAGCTTTCAGTACCTGCT 5916
 Qy 1908 GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLeuHisProGlyAsnIle 1927
 Db 5917 GGTCAAGATGACAGCTTCACTTGTGACCCCAAGAGCCCTCCCTCTGGCAACATC 5976
 Qy 1928 ProGlnSerGlyGlnAsnGlnLeuGlnProLeuLysProSerProSerSerAspAsn 1947
 Db 5977 CCAGAGTCCGGGCAATCAGTGTTCACGCCCTTAAGCCATCTCCCTCCAGTACAC 6036
 Qy 1948 LeuThrSerAlaPheThrSerAspGlyAlaIleSerValProSerLeuSerAlaProGly 1967
 Db 6037 CTCATATTAGCCTTACCAAGATATGTGTCATTTCAGTACCAAGCTTTCGCTCAGAT 6096
 Qy 1968 GlnGlyThrSerSerThrAsnThrValGlyAlaThrValAsnSerGlnAlaAlaGlnAla 1987
 Db 6097 CAAGGACCCAGCAGCAAACTCTGTGGGCAAGGTCAGCAAGCCGCCCAAGCT 6156
 Qy 1988 GlnProProAlaMetThrSerSerArgLysGlyThrPheThrAspLeuHisLysLeu 2007
 Db 6157 CAGCTCTGTCATACGTCAGCAGAGAGGCAATTCACAGATGACTGCAAGATTG 6216
 Qy 2008 ValAspAsnThrAlaArgAspAlaMetAsnLeuSerGlyArgArgGlySerLysGlyHis 2027
 Db 6217 GTAGACATTTGGGCGCCGAGATCCCATGATCTCTCAGAGAGAGAGAAAGGAGCAC 6276
 Qy 2028 MetAsnThrGlyGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuGlyIle 2047
 Db 6277 ATGAATTTATGAGGGCTTGGAAATGCAAGAAATTTCTTGACCTGGGGCACTGTGATC 6336

Qy 2048 SerMetThrSerAsnLeuGlyGlySerAlaProIleSerAlaAlaSerAlaThrSerLeu 2067
 Db 6337 TCCATGACCTCGAAGCTGGGTGCTCGCCCCATCTCTGACATCAGATCAGTACCTCTTA 6396
 Qy 2068 GlyHisPheThrLysSerMetCysProProGlnGlnThrGlyPheProAlaThrProPhe 2087
 Db 6397 GGTCACTTACCAAGTCTATGTCGCCCCCAACAGATGATGCTTCCAGCTACCCCATTT 6456
 Qy 2088 GlyAlaGlnThrSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlnPheGlnPro 2107
 Db 6457 GGGCTCATATGAGTGGGACGGGTGGCCGACACACAGCCTTGGCCAGTTCCAACT 6516
 Qy 2108 ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsn 2127
 Db 6517 GTGGAACTGCTCTCTTGCAGAAATTTCAATCAGCAATTTGGAGAAATCATCAGCAAC 6576
 Qy 2128 ProProGlySerAsnLeuArgThrThr 2136
 Db 6577 CCCCAGCTCCAACTGCGGACCACT 6603

RESULT 9

ACH03772 standard; DNA; 7065 BP.

ACH03772;

26-SEP-2003 (first entry)

DNA encoding novel human protein kinase #9.

Human; kinase; ds; gene; cosmetic application; nutraceutical application.

Homo sapiens.

US6541252-B1.

01-APR-2003.

14-MAY-2001; 2001US-00854856.

19-MAY-2000; 2000US-0206015P.

(LEXI-) LEXICON GENETICS INC.

Walke DW, Hillbun E, Donoho G, Turner CA;

WPI; 2003-575927/54.

P-PSDB; ABO44389.

New nucleic acid encoding novel human proteins, useful in cosmetic and nutraceutical applications.

Disclosure; Page; 11pp; English.

The invention relates to a new isolated nucleic acid encoding a novel human protein kinase. The nucleic acid is useful in cosmetic and nutraceutical applications. The present sequence represents DNA encoding a novel human protein kinase. Note: The sequence data for this patent did not from part of the printed specification but was obtained in electronic format directly from USPTO at

seqdata.uspto.gov/sequence.html?docID=6541252B1

Sequence 7065 BP; 1911 A; 1940 C; 1601 G; 1610 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 0 Length: 7065
 Score: 10513.00 Matches: 2108
 Percent Similarity: 88.50% Conservative: 0
 Best Local Similarity: 88.50% Mismatches: 0
 Query Match: 97.23% Indels: 274
 Gaps: 2

US-10-010-720-14 (1-2136) x ACH03772 (1-7065)

```

QY 1 MetSerGlyValAlaGluValGlnSerSerThrProGlySerLeuPheLeuSerPro 20
Db 1 ATGTCTGGCGGGCGCGCAGAGAGACAGACACACTCCCGGTTCTCTCTCTCGCGG 60
QY 21 ProAlaProAlaProAlaSerGlnSerSerSerSerSerValGlyGlyLeuSerGly 40
Db 61 CCGGCTCTCGCGCGCGCAGAGATGGCTCCAGCTCCGATTCCTCGTGGGGAGAACTCGGGA 120
QY 41 AlaAlaAlaAlaAlaAlaAlaValThrGlyValThrGlyValThrGlyValThrGly 60
Db 121 GCGCGGCGCGCGCGCAGAGCTGTGACCGCGCAGAGACCGAGGATACAGCGCGCGCGCACACT 180
QY 61 MetAspLysAspSerArgGlyValAlaAlaThrThrThrThrThrThrThrThrThrThr 80
Db 181 ATGGACAAAGAGACAGCGCTGGGGCGCGCGCGCACCTACACCACTGAGACCGCTTCTTC 240
QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
Db 241 CGCGCGAGCGCTCATCTGCGACTCCCAATGCACTGCACTGAGCTTCCGGGCTTCTCTT 300
QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
Db 301 TCCCTGCGCGCGCGCAGATCCCGCGGCTGTCCGCGAGAGTGTCCACCGAGCGCGCAC 360
QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
Db 361 CGGAGAGAGACCGTGAACCGCCACCGCCACTTCCAGGTAAGCCCGACAGCTTCCAGCCCT 420
QY 141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db 421 GCCCGCCCTGGGGGAAACAGCGCGCTGCGCGCGCTGCGCTGCACTGTCCCGAGAGTACC 480
QY 161 SerLysAspAspProValSerGlnProSerLeuValGlySerLysGluGluProProPro 180
Db 481 AGCAAAAGCCCGCGAGAGTCCAGCTTAGCTTGCGGGAGCAAAAGAGAGACCGCGCGG 540
QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGlnGluGluArgSerGlnGln 200
Db 541 GCGAGAGTGGCAGCGCGCGCGCGCGCGCCACAGAGACCAAGAGAGAGAGAGAGAGAGAG 600
QY 201 GlnAspAspIleGluGluLeuGlnThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
Db 601 CAGAGTATATCGAAGAGCTGAGAGCAAGCGCGTGGGAATGTCTAACAGATGGCGCTTT 660
QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
Db 661 CTCAAGTTTGACATCGAAATCGGAGAGGCTCTTTAAGACGGTCTCAAAAGGCTCTGAC 720
QY 241 ThrGluThrThrValGluValAlaAlaTrpCysGluLeuGlnAspArgLysLeuThrLysSer 260
Db 721 ACTGAAACCAACCGTGAAGTGGCTGTGTGATGCAATCGAGATCGAAATTTACAAAGTCT 780
QY 261 GluArgGluArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280
Db 781 GAGAGGAGAGATTTAAAGAAAGAGCTGAATGTTAAAGGTCTTCAGCATCCCAATATT 840
QY 281 ValArgPheTyrAspSerTrpGlnSerThrValLysGlyLysCysIleValLeuVal 300
Db 841 GTTAGATTTTATGATCTCTGGAAATCCACGATAAAGAGAAAGAGTGCATTGTTTGTG 900
QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Db 901 ACTGAACTTATGAGCTGTGGAACCTTAAACGATCTGAAAGAGTTTAAAGTATAG 960
QY 321 IleLysValLeuArgSerTrpCysArgGlnIleLysGlyLeuGlnPheLeuHisThr 340
Db 961 ATCAAGTTCTTAAGAGCTGTGCGTCAATCTTAAAGTCTTCAAGTTCTTCAATACT 1020
QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 1021 CGAACTCCACTTATCATTTACCGGAGATCTTAAATGTGACAAACATCTTTATCACCGGCTT 1080

```

```

QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
Db 1081 ACTGGCTCAGTCAAGATTTGAGAACCTCGGTGCGCAACCTTGAAAGCGGGCTTCTTTGGC 1140
QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
Db 1141 AAGAGTGTGATNGTACCCCGAGATTACGGCCCTCGAGATGTATGAGAGAAATATGAT 1200
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGlyTyr 420
Db 1201 GAATCCGTTGACGTTTATGCTTTTGATGTCATGCTTGAGATGGCTTACATCTGAATAT 1260
QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgValThrSerGlyValLys 440
Db 1261 CTTTACTGGAGTGCAAAATGCTGCACAGATCTACCGTGGGTACAGATGGGGTGAAG 1320
QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
Db 1321 CCAGCACTTTTGACAAAGTACATTCCTGAGTGAAGAGAAATATGAGAGATGCATA 1380
QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnIleAlaPhePheGln 480
Db 1381 CGACAAACCAAGATGAAAGATATTCATCAAGACCTTTGAAACCATGCTTCTTCCAA 1440
QY 481 GlnGluThrGlyValArgValGluLeuAlaGluLysAspGlyGluLysIleAlaIle 500
Db 1441 GAGGAAACAGAGTACGGGTGAAATGACAGAGAGATGTGAGAGAAAAATAGCCATA 1500
QY 501 LysLeuTrpLeuAlaGlyIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
Db 1501 AAATTTATGGCTACGATTTAGAGATTTTAAAGAAATTTAAAGGAAATACAAAGATATGAA 1560
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet 540
Db 1561 GCTATTGATTTCTTTGATTATTTAGAGAGATGTCCAGAGAGATGTTCACAGAAATG 1620
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 1621 GTTAGAGTCTGGGTATGCTGTGAAGGTGATCACAGACCAAGCGGTAAAGCTATCAAAAC 1680
QY 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGluGln 580
Db 1681 AGATATCATTTAATTAGAGAAACGAGAGCGGAGCTTGATCGGAGAGACAAAGAA 1740
QY 581 LysLysLysGlnGluGlnSerSerLeuLysGlnIleValGluGlnSerSerAlaSerGln 600
Db 1741 AAAAAAAGCAGAGAGAGAGAGCTCAAAACAGAGTAGAACAATCCAGTCTTCCAG 1800
QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1801 ACGAGAAATCAAGACGCTCCCTTGTGTCAGCACCGGCAATACCTTCTTACCACTTCA 1860
QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db 1861 GCTTCACTTTCTACAGATGAGAACTCGAAGAACCTGAGGACAGTCAACATCAACACTA 1920
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1921 CAGTACCAAGCAACCGATATATCTGTGTATCTGATGGAGCGGTTCACAGTGTCAAGGGA 1980
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1981 TCTCTGTCTTCAAGATCTCAGATGAGGAGCCCAAGACAGATTCATATGTGTCCCA 2040
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2041 CATGAACAGGCAATCTACAGGACAGTCCAGGGGATATACCTTCTACTGTCCAGCA 2100
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 2101 CAGTCTCAGCCCATGAGGATATATCCACCTCAAGTGTG----- 2139

```


4237 GTTGTTCAGTACAGCACTGTATCTTCAGTAACTTTCAGCACTTCAGCTCTGCA 4296
QY 1215 GLYISerThrAlaThrProGlyProIleProAlaValSerGlnGlnAla 1234
Db 4297 GGGGGCAGTACGTACACCCAGTCTTACAGCTCCAGCTGTAGTATCTAGAGAGAGCA 4356
QY 1235 GLYSerThrValGlyAlaThrLeuThrSerValSerThrThrSerPheProSer 1254
Db 4357 GGCAGACACTAGTGGAGCCACATTAACATGATTTTACACCACTTCATTCACAC 4416
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrThrProThrLeuAlaGlu 1274
Db 4417 ACAGCTTCAAGCTGTCCATTCAGCTTACAGCACTTCACTTCACTTCACTTCACTG 4476
QY 1275 ThrValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu 1294
Db 4477 ACCGGTAGTATGAGGCACTCCTAGTATGAGATTCATTCAGTACAGTCACTGATTTG 4536
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerProGlyAlaGlyValSerSerTy 1314
Db 4537 GCTTCTCCCTCTGCAACATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4596
QY 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4597 ATTCTCAGCTGTGGGCTGTGATCTTGTGTCATTCATTCATTCATTCATTCATTCCT 4656
QY 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlyValProSerIle 1354
Db 4657 ATTCTTCCCAAGAGCAAGCACTTCTTACACTTTTATTCCTCCCAAGTACTTCTCTCT 4716
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4717 CCACCTTGGTACAGCTGTGCTGCAATGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 4776
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisSerThrHisGlyProGlyValAspSer 1394
Db 4777 CAGCTTCAACAGCTTGTCTTCTCCACCAAGCCCACTTCTTCTTCTTCTTCTTCTTCT 4836
QY 1395 AspThrGlnProLysAlaProGlyIleAspAspIleLysThrLeuGlnGlyLysLeuArg 1414
Db 4837 GATACACAAACCAAGCTCTGGAATTCATGACATTAAGACCTTGAAGAAAGCTGCGG 4896
QY 1415 SerLeuPheSerGlnHisSerSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 4897 TCTCTGTTCAGTGAACACAGCTTCATCTGAGCTCAGCATGCTGCTGCTGCTGCTGAGACC 4956
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 4957 TCACCTAGTACATGAGAGCACTGTACACAGCAAGCACTCCCACTTCTGCTGTGCAACAGC 5016
QY 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 5017 AAACCTCCGACTTCTTACCAAGATACCTTACTTACCAACCAATTTTACCACTGAGAACCA 5076
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 5077 GTTGTCTTTCAGTACACAGCTGTGTCACACTGGGCAAGTTTCTTACCCCACTGAGCACT 5156
QY 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 5137 ACTACATCAGAGTGAAGAACTGTGAAGCTCTCCCAAGCACTTCAAAAGCTCCG 5196
QY 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 5197 GTGTGTCCAGTGGTACTGAACTTCCAGCAAGTACTTACCCAGCAAGCTGCCACT 5256
QY 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 5257 TTTCCAGAGACTTCTTCAACCACTCCAGCAACCTTCAAGAGATCTTGATGTCTCAATTG 5316
QY 1555 ArgArgThrLeuSerProGlnLysIleThrValThrSerAlaValGlyProValSerMet 1574
Db 5317 AGAAGAACCTTGTATCCAGATATATCAAGTACTTCTGTGGTGTCTGTCTCATG 5376
1575 AlaAlaProThrAlaAlaThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 5377 GGGGCTCCAAAGCAATCAAGAACAGAGAACAGCCCTCAGAAAGGTGTCTTCTCAAGTC 5436
QY 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysSerGlyArgPhe 1614
Db 5437 AAAGAAGCCCTGTCTTACCACTTGTTCAGAGCTGGTGTGTGTGTGTGTGTGTGTGTGT 5496
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
Db 5497 CAGTTTCTGTTCAGCAGCAGCGGTCCCAAGAAAGAGGGGAAAAATAGTCANAAGATGA 5556
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerSerValLeuSerSerSerPro 1654
Db 5557 AAGTGTGTTCATTTGAATTCAGCAACCTCAGAGTCTCCTATCAAGTATGATGTCCA 5616
QY 1655 GluSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 5617 GAGAGTACTTGTGTGAACCAAGCCCAATGGCATACATCTCTGTGTATCTTCTCAGAT 5676
QY 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnPro 1694
Db 5677 GTGCCAGAGAGTCCCAAAACTACTGCTCAGAGGCAAGTCAACACTGGCAAGCT 5736
QY 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 5737 ACCAAGGTGAGCGTTTCAGGTGACCACTACAGCAACCAAGTGGTGTCTCTCTCTCTCT 5796
QY 1715 SerLysThrGlnAspLysIleThrAspThrLysLysGlnGlyProValAlaSerProPro 1734
Db 5797 TCAAAACGTAGAGCAAGATACCTGACACAAAGAAAGAGAACAGTGGCATCTCTCTCT 5856
QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlnLysProGln 1754
Db 5857 TTTATGATTTGGAAACAAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5916
QY 1755 LeuSerGlnProSerHisLysAsnGlyProSerSerAspProGlnAlaAlaPheLeuSer 1774
Db 5917 CTGTCAAGCTTTCACATCTTAAATGGGCGCTTCTGTGACCCGAGGCGCTTTTATAGT 5976
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
Db 5977 AGGATGTGATGATGTGTTCCGTAGTCCACACTGCCCCATCAGCTCAGCTCAAAAGCC 6036
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSer 1814
Db 6037 CTTCCTAGCCAGATCTTAACTTAAAGCTTAACTTAACTTAACTTAACTTAACTTAACT 6096
QY 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuGlnLeuArgArgLeuArg 1834
Db 6097 AGCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6156
QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
Db 6157 GATTAACATCTCAAAAGATTTCAAGACTTCGACAGTGCACAGAGCATGAAATTAATCT 6216
QY 1855 LeuThrThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu 1874
Db 6217 TTTGATACCAAACTGGGCAAGGTGCCCCCTGTGTATTTATTTCCCCCAGCTGCCCTT 6276
QY 1875 SerGlyArgArgArgArgProThrThrLysSerLysGlySerLysSerSerArgSerSer 1894
Db 6277 TCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6336
QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
Db 6337 TTGGGGAATAAAGCCCCAGCTTTCAGTATCTGTCTGTCTGTCTGTCTGTCTGTCTGT 6396
QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
Db 6397 TTGCACCCCCAGCAGACCTCCACCCCTGTGGAACAATCCCAAGTCCGGGAGAAATCAG 6456

QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
 PT CTGTTAAGAGCCCTTAAGCATCTCCCTCCAGTACAACTCTATTGAGCTTCACAGT 6516
 Db 6457 CTGTTAAGAGCCCTTAAGCATCTCCCTCCAGTACAACTCTATTGAGCTTCACAGT 6516
 QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
 Db 6517 GATGATGCAATTCATGACCAAGAGCTTCTGCTCCAGGTCAAGGAAACACAGCAGCAAAAC 6576
 QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaGlnAlaGlnProProAlaMetThrSer 1994
 Db 6577 ACTGTTGGGGCAAGGAAAGCAAGCAAGCCGCAAGCTCAAGCTCTGCGCATGACCTCC 6636
 QY 1995 SerArgGlySerGlyThrPheThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAsp 2014
 Db 6637 AGCAGGAAGAGGACATTCACAGATGACTTGCACAAAGTTGGTAGCAATTGGGCCCGAGAT 6696
 QY 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnLysGlyGlyProGly 2034
 Db 6697 GCCATGAATCTCTCAGAGCAGAGAGAGAAAGCAAGGCAATGATTTATGAGGGCCCTGGA 6756
 QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuLysLysSerMetThrSerAsnLeuGly 2054
 Db 6757 ATGGCAAGGAAGTCTCTGCACTGGGCAACTGTGCATCTCATGACCTCGAACTG3GT 6816
 QY 2055 GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074
 Db 6817 GGCTCTGCCCCCATCTCTGCAGCATCAGCTACCTCTAGTCACTTCAACCAATCATG 6876
 QY 2075 CysProProGlnGlnLysGlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThr 2094
 Db 6877 TGCCCCCAGCAGCATGATGGCTTTCAGCTACCCCATTTGGCGCTCAATGAGTGG3AAG 6936
 QY 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
 Db 6937 GGTGGCCAGCAGCAGCATCTGGCCAGTTCCAACTGTGG3AAGCTGCTCTTGAG 6996
 QY 2115 AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg 2134
 Db 6997 AATTTCACATCAGCAATTTCAGAAATCCATCAGCAACCCCGCAGCTCCAACTGG3G 7056
 QY 2135 ThrThr 2136
 Db 7057 ACCACT 7062

RESULT 10
 ACH03786
 ID ACH03786 standard; DNA; 6231 BP.
 XX
 AC ACH03786;
 XX
 DT 26-SEP-2003 (first entry)
 XX
 DE DNA encoding novel human protein kinase #23.
 XX
 KW Human; kinase; ds; gene; cosmetic application; nutraceutical application.
 XX
 OS Homo sapiens.
 XX
 PN US6541252-B1.
 XX
 PD 01-APR-2003.
 XX
 PF 14-MAY-2001; 2001US-00854856.
 XX
 PR 19-MAY-2000; 2000US-0206015P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hilbun E, Donoho G, Turner CA;
 XX
 DR WPI, 2003-575927/54.
 XX
 P-PSDB; ABO44403.

PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutraceutical applications.
 XX
 XX Disclosure; Page; 11pp; English.
 PS
 XX The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents DNA encoding
 CC a novel human protein kinase. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=6541252B1
 CC
 XX
 SQ Sequence 6231 BP; 1750 A; 1657 C; 1412 G; 1410 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	6231
Score:	10508.00	Conservative:	2076
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	97.19%	Gaps:	0

US-10-010-720-14 (1-2136) x ACH03786 (1-6231)

QY 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrGlnHisArgPhePhe 80
 Db 1 ATGGCAAGGAAGAGCGGTGGGGGGCGCGCGCACTACCAACCACTGAGCAACCGCTTCTTC 60
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLysProLeu 100
 Db 61 CGCGGAGCGCTCATCTCGACTCCCAATGCCATGCGACCTGAGGCTTCCCGGCTTCTTCTT 120
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
 Db 121 TCCCTGCCCCAGCCACATCTCCGCGCTGTCCGAGACTGCTTCAACCG3AGCCCAAC 180
 QY 121 ArgGlnGlnThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 181 CGGAAAGAGACCGGAGCGCGCCAGCCCACTTCCAGGTAGCCGAGCGCTCCAGCGCT 240
 QY 141 AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 241 GCCGCCCTTGGGAAAGCGCGCTGCGCGCGCTGCGCGCTGCGCTGCGCTGCGCGCTG 300
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGlnProProPro 180
 Db 301 AGCAAGACCGCCAGAGTCTCCAGCTTGTGGAGACAAAGAGACCGCCGCCG 360
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGlnProGlnGlnGlnArgSerGlnGln 200
 Db 361 GCGAAGAGTGGCAGCGGGCGCGCGCGCGCAAGAGCCAGAGGAAAGGAGCGAGCAG 420
 QY 201 GlnAspAspIleGlnGlnLeuGlnThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
 Db 421 CAGATGATATCGAAGAGCTGAGACCAAGCCCGTGGAAAGTCTAACAGATGGCGCTT 480
 QY 221 LeuLysPheAspIleGlnIleGlyArgGlySerPheLysThrValLysGlyLeuAsp 240
 Db 481 CTCAAGTTGACATCGAATTCGCGAGAGGCTCTTTAAGAGGCTTACAAAGGCTGAGAC 540
 QY 241 ThrGlnThrThrValGlnValAlaTrpCysGlnLeuGlnAspArgLysLeuThrLysSer 260
 Db 541 ACTGAACACACCGTGAAGTGCCTCGTGTGAATGACAGATCGAAATTAACAAAGCT 600
 QY 261 GlnArgGlnArgPheLysGlnGlnAlaGlnMetLeuLysGlyLeuGlnHisProAsnIle 280
 Db 601 GAGAGGCGAGATTTAAAGAAAGCTGAAATGTTAAAGGCTTCAACATCCCAATAT 660
 QY 281 ValArgPheLysArgSerTrpGlyLysThrValLysGlyLysCysIleValLeuVal 300
 Db 661 GTTAGATTATTATGATTCTCGGAATCCACATGAAGAAAGAAAGTGAATGATTGTTG 720

QY 301 ThrGluMetThrSerGlyThrLeuThrTyLeuLeuArgPheLeuValMetLys 320
Db 721 ACTGAACCTTATGACGCTCTGGAAACACTTAAAGTATCTGAAAGGTTTAAAGTGAAG 780
QY 321 IleValLeuArgSerThrPyrArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
Db 781 ATCAAAAGTTCAAGAGCTGGTGGCCGTGAGATCTTAAAGGCTTCAGTTCTTCATACT 840
QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 841 CGAATCTCACCTTATCATTCATCCGCGATCTTAAATGTGACACATCTTATCATCCGCGCT 900
QY 361 ThrGlySerValLeuLysIleGlyAspLeuGlyLeuValThrLeuLysArgAlaSerPheAla 380
Db 901 ACTGCTCAGTCATGATGAGATTCGAGACCTCGGCTGGCAACCTGAAACGCGGCTCTTTGGC 960
QY 381 LysSerValIleGlyThrProGlnPheMetAlaProGlnMetTyrGlnGlnLysTyrAsp 400
Db 961 AAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlnTyr 420
Db 1021 GAATCCGTTGACGTTTATGCTTTGGGATGTCATGCTTGAATGGCTTACATTCGAAAT 1080
QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgValThrSerGlyValLys 440
Db 1081 CCTTACTGGAGGTGCAAAATGCTGCAACAGATCTACCTCGGCTGACAGTGGGCTTGAAG 1140
QY 441 ProAlaSerPheAspValAlaIleProGlnValLysGlnIleIleGlnGlyCysIle 460
Db 1141 CCAGCCAGTTTGGACAAAGTACATCTTGAAGGAAATTTATGAAAGGATGCACTA 1200
QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
Db 1201 CGACAAACAAAGATGAAAGATATTCATCAAAAGCCTTTGAACCATGCTCTTCCAA 1260
QY 481 GlnGlnThrGlyValArgValGlnLeuAlaGlnGluAspAspGlyGlyLysIleAlaIle 500
Db 1261 GAGAAACAGAGAGTACGGGTGAAATTGACAAAGAGATGATGAGAAATTAATGACCTA 1320
QY 501 LysLeuThrPheLeuArgIleGluAspIleLysLeuLysGlyLysTyrLysAspAsnGlu 520
Db 1321 AAATTTATGCTGATGATTTGAAGATTTAAAGAAATTTAAAGGAAATCAAAAGATTAAGAA 1380
QY 521 AlaIleGlnPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGlnMet 540
Db 1381 GCATATGAGTTCTTTTGTGATTTAGAGAGAGATGCTCCAGAAAGATGTTGCAACAAATG 1440
QY 541 ValGlnSerGlyTyrValCysGlnGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 1441 GTACAGCTGGGATGCTGTGAGAGGTGATCAAAAGCATGCTTAAAGCTATCAAAAGAC 1500
QY 561 ArgValSerLeuIleLysArgLysArgGlnGlnArgGlnLeuValArgGlnGlnGlnGlu 580
Db 1501 AGAGTATCATTTAATTAAAGAGAAACGAGAGCGGATGTTGATGAGAGAGACAAAGAA 1560
QY 581 LysLysLysGlnGlnGlnGlnSerSerLeuLysGlnGlnValGlnGlnSerSerAlaSerGln 600
Db 1561 AAAAAGAGAGAGAGAGAGAGAGCTCAACAGAGAGTAAAGACATCAAGCTCTCCAG 1620
QY 601 ThrGlyLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1621 ACGAGATCAAGACGCTCCCTCTGCTAGACACCGGCTACCTACTGCTTACCACTTCA 1680
QY 621 AlaSerValSerThrGlnValGlnProGlnGlnProGlnAlaAspGlnHisGlnGlnLeu 640
Db 1681 GCTTCACTTCTACCAAGTAACTGAAAGAACTGAAAGAACTGAAAGAACTGAAAGAACT 1740
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1741 CAGTACCAAGAACCCAGATATATCTGTTATCTGATGAGACGTTGACAGTGTGCAAGGA 1800
QY 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1801 TCCTCTGCTTACCAAGATCTGAGTGAACACCAACAGAGTTTCATATGATGCTCCAA 1860
QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlnHisIleProSerThrValGlnAla 700
Db 1861 CATGAACAGGACATCTTACAGGACAGACAGTCCAGAGGATATACCTTCTACTGTCAGGA 1920
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 1921 CAGTCTACGCCCCAAGGGGATATATCCACCTCAAGTGTGGCACAGGGGACAGACAGGCT 1980
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 1981 CAGCATCTCTCAAGTACCTTAAAGAGGATTCATCTCCCAACCCATACATCATCTCCAG 2040
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2041 CAGCAGCAGGAGATACAGACAGACAGCCCTCCTCAACAGACAGTCAAGTATTCATTTCA 2100
QY 761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2101 CAGACATCACTTCAGTGAAGCCACTGACCTGACAGCCAGTGAAGTCAAGCTTCMACTCCA 2160
QY 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
Db 2161 CAAGCTTGCCTCAAGTATCAAGCTGGAAGAAACAGAGTATCTCAGGAGTCTTCAAGTCTC 2220
QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db 2221 CCTGACAGGCGAGTTGACAGTACAGACAGCCCAAGCTACCCAGCCGACCACTTGGCTTCC 2280
QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyLysGlnAsnVal 840
Db 2281 TCTGTAACAGTGAACATTCAGATGCTTCAAGTATGAGTATGAGTATGAGTATGAGTATG 2340
QY 841 ProSerSerSerGlyArgHisGlnGlyArgThrThrLysArgHisTyrArgLysSerVal 860
Db 2341 CCATCTTCCAGTGAAGGATGAGAGGAAAGAACTCAAAACCGCATTCACGAAATCTGTA 2400
QY 861 ArgSerArgSerArgHisGlnLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
Db 2401 AGGAGTGCCTTCCAGATGAAATAAACTTCAAGCCCAAAATTAAGAAATTTGAATGTTCA 2460
QY 881 AsnLysGlyAspArgValValGlnCysGlnLeuGlnThrHisAsnArgLysMetValThr 900
Db 2461 AATTAAGAGAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
QY 901 PheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAlaThrIleMetValAsnAsn 920
Db 2521 TTCAAAATTTGACCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580
QY 921 AspPheIleLeuAlaIleGlnArgGlnSerPheValAspGlnValArgGlnIleIleGln 940
Db 2581 GACTTATTTCTAGCAATGAGAGAGAGTCTGTTGTGATCAAGTGCAGAAATTAATTGAA 2640
QY 941 LysAlaAspGlnMetLeuSerGlnLysAspValSerValGlnProGlnGlnLysGlnGln 960
Db 2641 AAAGCTGATAAATGCTCAAGTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 2700
QY 961 GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGlnGlyGln 980
Db 2701 GAGAGCTTACAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
Db 2761 TTCAAAACAAACAAATCTCTGCTTCCATGCTCCACAGCAATTAAGCATTCCTACACGTTCT 2820
QY 1001 LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGlnSer 1020
Db 2821 TTAACCTCAAGTGTTCATTTGCGGAGAGCGGTTTATATGATGATGATGATGATGATGAT 2880
QY 1021 ArgLeuArgGlnSerLysValPheProSerGlnIleThrAspThrValAlaAlaSerThr 1040

Db 2881 CGATTACGAGATCAAAAGTTTCCCAAGTGAATTAACATACAGTTGCTGCTTACA 2940
QY 1041 AlagInserProglIymeAsnLeuSerHisSerIaSerSerLeuSerLeuGlnIalA 1060
Db 2941 GCTAGAGCCCTGGAAATCACTTGTCTCACTGCTGATCATCTTATGCTTCAACAGGCC 3000
QY 1061 PheSerGluLeuAgtAgtAglAmeThrGluGlyProAsnThrAlaProProAsnPh 1080
Db 3001 TTTTCTGACTTACACGCGCCCAATATGACAGAGACCAATACAGACCTCCAAACTTT 3060
QY 1081 SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAglIyVal 1100
Db 3061 AGTATACAGAGCAACATTTCCAGTAGTACCTCTTCTTAAGTAGAGTTGCTGAGTC 3120
QY 1101 ProThrThrAlaAlaIaThrAlaProValProAlaThSerSerProProAsnAspIle 1120
Db 3121 CCAACACAGACAGACCCACAGACACAGTCCCTGCAACAGACGCCCTCTTAAGACATT 3180
QY 1121 SerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAglIyVal 1140
Db 3181 TCCACATCAGTAATTCAGTCTGAGGTTACAGTCCCACTGAAGAGGAGATTGCTGAGTT 3240
QY 1141 AlathSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer 1160
Db 3241 GCCACACAGACAGAGTGTGTACTTCAAGTGTCTCCCATACCACTGTGTCTGAATCA 3300
QY 1161 ProValLeuSerSerValIaSerSerIleThrIleProAlaValIaSerIleSerThr 1180
Db 3301 CCACTACTTTCCAGCGTAGTTTCAAGTATCAATACCTGCAAGTTGTCTCAATATCTACT 3360
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValIaSerSerThrAla 1200
Db 3361 ACATCCCGCTCACTTCAAGTCCCAACATCACTGAGATCGTTGTTCTAGTACACA 3420
QY 1201 LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlyIySerThrAlaThr 1220
Db 3421 CTGATCCTTCAAGAACAGTTCAGCACTTCAAGCTCTGCAAGGAGGAGTACCTGACCC 3480
QY 1221 ProGlyProLysProProAlaValIaSerGlnIalAglIySerThrThrValGly 1240
Db 3481 CCAAGTCTTAAGCTCCAGCTGTAGTATCTCAAGACAGACAGGACAGACCTACTGTGGA 3540
QY 1241 AlathLeuThrSerSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 3541 GCCAATTAACATCAGTTTCTCAACACCACTTCACTCCCAAGACAGCTTCAAGCTGCC 3600
QY 1261 IleGlnLeuSerSerSerThrThrProThrLeuAlaGluThrValIaValSerAla 1280
Db 3601 ATTCACTTACAGACAGTACTTCTACTCTTACCTTGAACCCGTGTGTAGTTAGCCCA 3660
QY 1281 HisSerLeuAspIyThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 3661 CACTACATGATTAAGACATCTCAAGACGTACAACTGATGAGCTTCTCCCTCTCTGCA 3720
QY 1301 ProSerSerSerSerSerProGlyIaGlyIaSerSerIyIleSerGlnProGlyIy 1320
Db 3721 CCATCTTCTCTTCTCTCTCTGAGACAGAGAGTCTAGTATATTTCCAGCTGCTGAGG 3780
QY 1321 LeuHisProLeuValIleProSerValIleAaSerThrProIleLeuProGlnAlaIa 1340
Db 3781 CTGACATCTTGTGCTATTCATCACTGATGAGTCTTACTCTTATCTTCCCAACACACA 3840
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db 3841 GGAACCTACTTCAACCTTATTAACCAAGTACCTAGTATCCCACTTGTGTAAGCCT 3900
QY 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 3901 GTTGCCATAGTGCCTGCTGTAACAGACACATTAATCAATGACAGCTCAACAGCTTTG 3960
QY 1381 LeuProAsnGlnProHisThrHisGlyProGluValAspSerAspThrGlnProIyAla 1400
Db 3961 CTTCCCAACCAAGCCCATATCTATGTCTGTGAAGTAGATTCGATACCAACCAAGCT 4020

QY 1401 ProGlyIleAspAspIleIyThrLeuGlnGluIyLeuArgSerLeuPheSerGluHis 1420
Db 4021 CTTGATATGATGACATTAAGACCTTGAAGAAAGCTGCGGTCTGTTCACTGAACAC 4080
QY 1421 SerSerSerGlyIaGlnHisAlaSerValSerLeuGluThrSerLeuValIleGluSer 1440
Db 4081 AGCTCATCTGAGTCCAGATGCTCTGTCTCATCTGAGACCTTCACTAGTACATGAGAGC 4140
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerIyLeuLeuThrSerThr 1460
Db 4201 ACTGTCAACAGGACATCCCACTACTGCTGTGCAACCAAGAACTCTTACTTCTTACC 4200
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4201 ACAAGTACTTGCTTACCAACCAACCAATTTACATGAGACAGTGTGCTTCCAGTTACA 4260
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValIys 1500
Db 4261 CCAGTGTCAACCTGAGCAAGTTTCTAACCCAGTCACTACTTACATCAGAGTGA 4320
QY 1501 ProGlyThrAlaProSerIyProProLeuThrIyAlaProValLeuProValGlyThr 1520
Db 4321 CTTGAACTGCTCTCTTCAAGCCACCTTAACTTAAGCTCCGCTGCTGCCAGTGGTACT 4380
QY 1521 GluLeuProAlaGlyThrLeuProSerGluGlnLeuProPheProGlyProSerIyLeu 1540
Db 4381 GAATCTTCAGAGGATCTTACCCAGCAGAGCTGCCACTTTCCAGACCTTCTCTA 4440
QY 1541 ThrGlnSerGlnProLeuGluAspLeuAspAlaGlnLeuAgtAgtThrLeuSerPro 1560
Db 4441 ACCAGTCCCAAGCAACCTTGAAGATCTTATCTCAATTAAGAAAGAACTTATGCTCA 4500
QY 1561 Glu**IleThrValIleThrSerAlaValIaGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4501 GAGATKATCAGAGTACTTCTGCGGTGTGCTTGTCTTCAAGTGGCGGCTCCAAACAGCATC 4560
QY 1581 ThrGluIaGlyThrGlnProGlnIyIyValSerGlnValIyGluIyProValIleu 1600
Db 4561 ACAGAGAGAGAACACAGCTCAGAGAGGTCTTCTCAAGTCAAAAGAGGCCCTGCTCTA 4620
QY 1601 AlathSerSerGlyIaGlyIaPheIyMetGlyIyArgPheGlnValSerValAlaIaIa 1620
Db 4621 GCAACTAGTTCAGAGCTGCTGTTTAAAGTGGACATTTCAAGTTCTGTTGACACA 4680
QY 1621 AspGlyIaGlnIyGluGlyIyAsnIySerGluAspAlaIySerValHisPheGlu 1640
Db 4681 GACGTGCCCAAGAAAGAGGTAATAATGACAGAAATGCAAGTCTGTTCAITTTGAA 4740
QY 1641 SerSerThrSerGluSerSerValLeuSerSerSerProGluSerThrLeuValIys 1660
Db 4741 TCCAGACCTCAGAGTCTCTCACTGCTATCAAGTAGTACTCCAGAGATCCTGGTGA 4800
QY 1661 ProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGluSerAlaHis 1680
Db 4801 CCAGAGCCGAATGCAATCACTCCCTGATCTTCAAGTGTGCCAAGAGTCCAC 4860
QY 1681 LysThrThrAlaSerGluAlaIySerAspThrGlyGlnProThrIyValIaGlyArgPhe 1700
Db 4861 AAACCTACTGCTCAGAGGCAAGTCAAGACCTGGCAGCTTACCAAGTGTGAGCTTT 4920
QY 1701 GlnValThrThrThrAlaAsnIyValIaIyArgPheSerValSerIyThrGluAspIy 1720
Db 4921 CAGGTGACATCAGCAACAAAGAGTGGTGTCTCTGTATCAAAAGTGAAGACAAG 4980
QY 1721 IleThrAspThrIyIyGlyGlyIyProValAlaSerProPheMetAspLeuGluGln 1740
Db 4981 ATCACTGACACAAAGAAAGAGACAGTGGATCTCTCTTATATGATTTGAACAA 5040
QY 1741 AlaValLeuProAlaValIleProIyIyGluIyProGluLeuSerGluProSerHis 1760
Db 5041 GCTGTTCTTCTGCTGTATACCAAGAAAGAAAGAGGCTGAAGCTCAGAGCTTCAAT 5100

QY	1761	LeuAnngLYProSerSerSerpProGluAlaAPheLeuSerAArgAspValAspAspGly	1780
Db	5101	CTAAATGGGCGCCTCTTCGACCCGGAGGCCGCTTTTAAAGTAGGGAAGTGGATGATGCT	5160
QY	1781	SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu	1800
Db	5161	TCCGGTAGTCACACCTCGGCCCATAGCTGAGCTGAAGAGCCTTCTAGCCAGAACTTA	5220
QY	1801	SerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSerSerAspAsnGluSerAsp	1820
Db	5221	AGTCAAAGCCTTAGTAATTCATTAACTCTCTTCATGAGTAGAGCAATATGAGTCAGAT	5280
QY	1821	ILegIAuSPGLuAspLeuLysLeuGluLeuArgAArgLeuAArgAspLysHisLeuLysGlu	1840
Db	5281	ATCGAAGATGAAGACTTAAAGTTAAGTGCACACACATCAAGATAACATCTCAAAAG	5340
QY	1841	ILegIAuSPLeuGlnSerAArgGlnLysHisGlnILegIuSerLeuTyThLysLeuGly	1860
Db	5341	ATTGAGAGACTGCAGAGTGCACAGAGATGAATTTGAATCTTTGTATACAAACTGGGC	5400
QY	1861	LysValProProAlaValILeIleProProAlaAlaProLeuSerGlyAArgAArgArg	1880
Db	5401	AAGGAGCCCCCTGCTGTTATTAATCCCCAGCTGCTCCCTTCAGGAGAGAAGCAGCA	5460
QY	1881	ProThLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerPro	1900
Db	5461	CCCACTTAAGCAAGCAAGCAAACTTAAGTGAAGCACTTCTTGAGGGAATAAAGCCCC	5520
QY	1901	GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr	1920
Db	5521	CAGCTTTCAGGTAACCTGCTGCTGTCAGAGTGCAGCTTCAGCTTGACCCCCAGCAAGCC	5580
QY	1921	LeuHisProProGlyAsnILeProGluSerGlyGlnAsnGlnLeuGlnProLeuLys	1940
Db	5581	CTCCACCCCTCGGCAACATCCCAAGTCCGGGCAAGATCAGCTGTATACAGCCCTTAAG	5640
QY	1941	ProSerProSerSerAspAsnLeuTySerAlaPheThrSerAspGlyAlaIleSerVal	1960
Db	5641	CAATCTCCCTCAGAGACACCTCTATTCACCTTCACCACTGATGAGGCCATTTCCAGTA	5700
QY	1961	ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal	1980
Db	5701	CCAAAGCCTTCTGCTCCAGTCCAGAGAACGACGACACAACTGTTGGGGCAACAGTG	5760
QY	1981	AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerAArgLysGlyThrPhe	2000
Db	5761	AACAGCCAAAGCCGCCAAAGCTCAGCTCTCTCCATGACGTCCACACAGAAAGGCACTTC	5820
QY	2001	ThirAspAspLeuHisLysLysLeuValAspAsnThrAlaArgAspAlaMetAsnLeuSerGly	2020
Db	5821	ACAATATGACTTGACAAGTTGTGTGACAAATTTGGGCCGAGAGTGCATGAATCTCTCAGGC	5880
QY	2021	ArgArgGlySerLysGlyHisMetAsnTyArgLysGlyProGlyMetAlaArgLysPheSer	2040
Db	5881	AGGAGAGGAACCAAGGACACATGAATTTATGAGGGCCCTGGAAATGGCAAGGAATTTCT	5940
QY	2041	AlaProGlyGlnLeuGlyHisSerMetThrSerAsnLeuGlyGlySerAlaProIleSer	2060
Db	5941	GCACCTGGGCAACTGTGATCTTCATGACCTCGAAGCTCGGAGTGGCTTGCCCCCACTCT	6000
QY	2061	AlaAlaSerAlaThrThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnTy	2080
Db	6001	GCACACATCAGTACTCTCTTAGAGCATTCAACAAAGTCTATGTGCCCCCAACAGAGTAT	6060
QY	2081	GlyPheProAlaThrProPheGlyAlaGlnThrSerGlyThrArgLysProAlaProGln	2100
Db	6061	GGCTTTCAGTACCCTCAATTTGGGCTCAATGAGTGGAGCGGGTGGCCCAAGCAACAG	6120
QY	2101	ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn	2120
Db	6121	CCACTTGGCCAGTTCCAACTGTGGGAAGTGGCTCTCTTGAGATTTCAACATCACAGCAAT	6180
QY	2121	LeuGlnLysSerIleSerAsnProProGlySerAsnLeuArgThrThr	2136

[illegible]

QY 141 AlaAlaProGluGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db 241 GCCGCCCTGGGGAAACAGCCGTCGCGGGCCCTGCCCTCGACTGTCCTCCAGAGTACC 300
QY 161 SerIysAspArgProValSerGlnProSerLeuValGlySerIysGluGluProProPro 180
Db 301 AGCAAAACCCGCCAGGTGCCAGCTTGTGGGAGCAAAAGAGAGAGAGAGAGAGAGAGAG 360
QY 181 AlaArgSerGlySerGlyGlyGlySerAlaIysGluProGlnGluGluGluGluGluGlu 200
Db 361 GCGAAGATGGACGCGCGCGCGAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 201 GluAspAspIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 220
Db 421 CAGAGATGATTCGAAGAGCTGGAGACCAAGGCCCGTGGAGATGCTTAACGATGCCGCTTT 480
QY 221 LeuIysPheAspIleGluIleGlyArgGlySerPheIysThrValIleIysGlyLeuAsp 240
Db 481 CTCAAGTTTGCATCGAATCGCAAGGCTCTTTAAGACGCTCTCAAGGCTCTCAAGGCTCT 540
QY 241 ThrGluThrThrValGluValAlaIleProGluGluGluGluGluGluGluGluGluGlu 260
Db 541 ACTGAACACCGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 261 GluArgGluArgPheIysGluGluGluGluGluGluGluGluGluGluGluGluGlu 280
Db 601 GAGAGGACGAGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 281 ValArgPheIysArgPheSerThrGluSerThrValIysGlyIysGlyIysGlyIysGly 300
Db 661 GTTAGATTTTATGATTTCTCTGGGAAATCCACAGTAAAGAGAGAGAGAGAGAGAGAG 720
QY 301 ThrGluGluMetThrSerGlyThrLeuIysThrIleuIysArgPheIysValMetIys 320
Db 721 ACTGAACCTTATGACGCTCGAAGCACTTAAAGAGATCGAAGAGAGAGAGAGAGAGAG 780
QY 321 IleIysValIleuArgSerThrProCysArgGlnIleuIysGlyLeuGlnPheLeuHisThr 340
Db 781 ATCAAGTCTTAAGAGAGCTGGTGGCGGTGAGATCTTAAAGGCTCTTCACTTCTCATACT 840
QY 341 ArgThrProProIleIleHisArgAspLeuIysCysAspAsnIlePheIleThrGlyPro 360
Db 841 CGAATCCACCTATCATCATCACCGCATCTTAATGTGCAACATCTTATCACCGGCGCT 900
QY 361 ThrGlySerValIysIleGlyAspLeuGlyLeuAlaThrLeuIysArgAlaSerPheAla 380
Db 901 ACTGGCTAGTCAGATTCGAGAGACTCGGCTCGGAGACCTCGAGAGCGGGCTTCTTTGCC 960
QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetIysGluGluIysIysAsp 400
Db 961 AAGAGTGTGATAGTACCCAGAGTTCATGGCCCTGAGATGTATGAGAGAGAAATATGAT 1020
QY 401 GluSerValAspValIleValAlaPheGlyMetCysMetLeuGluMetAlaThrSerGlyIys 420
Db 1021 GAATCCGTTGACGTTATGCTTTGGAGTGTGATGCTTGAGATGGCTACATCTCAAAAT 1080
QY 421 ProIysSerGluCysGlnAsnAlaIleGlnIleIysArgArgValIleThrSerGlyValIys 440
Db 1081 CTTTACTGAGGTGCAAAAGTCTGACAGATCTACCGTGGCTGACCAAGTGGGGTGAAG 1140
QY 441 ProAlaSerPheAspIysValAlaIleProGluValIysGluIleIleGluGlyCysIle 460
Db 1141 CCAGCAGATTTGACAAAGTATGACATCTCTGAAGAGAGAGAGAGAGAGAGAGAGAGAT 1200
QY 461 ArgGluAsnIysAspGluArgIysSerIleIysAspLeuLeuAsnHisAlaPhePheGln 480
Db 1201 CGACAAACCAAGATGAAAGATATTCATCAAAAGACCTTTGAACCATGCCCTTCTTCAA 1260
QY 481 GlnGluThrGluValArgValGluLeuAlaGluGluGluGluGluGluGluGluGluGlu 500
Db 1261 GAGGAAACGAGGTACGGGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1320
QY 501 LysLeuTrpLeuArgIleGluAspIleIysIysIysLeuIysGlyIysIysIysAspAsnGlu 520

Db 1321 AAATTATGCTACGATATTGAAGATATTAAAGGAAATTAAGGAAATTAAGATATGA 1380
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluPheValAlaGlnGluMet 540
Db 1381 GCTATTGACTTTTCTTTGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 541 ValGluSerGlyIleValCysGluGlyAspHisIysThrMetAlaIysAlaIleIysAsp 560
Db 1441 GTAGAGTCTGGATATGCTGTGAAGGTGATCACAGACCATGCTTAAGCTATCAAGAGAC 1500
QY 561 ArgValSerLeuIleIysArgIysArgGluGluGluGluGluGluGluGluGluGlu 580
Db 1501 AGAGTATCATTAATTAAAGAGAAACGAGACAGGAGAGAGAGAGAGAGAGAGAGAGAA 1560
QY 581 LysIysIysGlnGluGluSerSerLeuIysGlnGlnValGluGlnSerAlaSerGln 600
Db 1561 AAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 601 ThrGlyIleIysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrSer 620
Db 1621 ACAGAGATCAAGAGAGCTCCCTTCTGAGACCGGACATACCTACTGCTTACCACTTCA 1680
QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlu 640
Db 1681 GCTTCAGTTCTTACCAAGATAGAACTGAAAGACCTGAGGAGATCAACATCAACACTA 1740
QY 641 GlnTrpGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1741 CAGTACAGACCAACCAATATCTGCTGTATCTGATGGAGAGGTGACAGAGGTCAAGGA 1800
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerThrGlySerGln 680
Db 1801 TCCTCTGCTTCAAGAGATCTCGATGAGGAGCCAAAGACAGTTTCATATGTTCCCAA 1860
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 1861 CATGAACAGGACATCTTACAGGAGCAGCAGCTCCAGGAGATATACCTTCACTGCTCAAC 1920
QY 701 GlnSerGlnProHisGlyValIleProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 1921 CAGTCTACGCCCTAGGGGTATATCCACCTTCAGAGTGGGAGAGGAGAGAGAGAGAGGT 1980
QY 721 GlnProSerSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 1981 CAGCATCTCTCAAGAGCTTAAAGGGGTTTCATCTTCCCAACCAATACATCTCTCAG 2040
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnIysLeuSer 760
Db 2041 CAGCAGCAGGAGAAATACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2101 CAGACATTAACCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
QY 781 GlnValLeuProGlnValSerAlaGlyIysGln 791
Db 2161 CAAGTCTGCTCAAGATATCAGTGAAGAAACAGGGCTTCCCACTCGAGTGCACAGAG 2220
QY 791 791
Db 2221 TACCCAGAGATTCAAATATTTGCTCCCTTCCAAAGTGGCTTCTGTTGATCCATCT 2280
QY 791 791
Db 2281 ACACTTACCCCTCCATGCGCAGACAGAGATCTGCTACACCTGGGTATCTTCCACA 2340
QY 791 791
Db 2341 GTGTGAGCCTTATGTGAATCAAACTTTTATGCTTATGAGGTGTGTAGAGAGAGAG 2400
QY 791 791

Db 2401 GTTCAGTGTCCACGACGAGGAGTTTACCAACAGCCCACTACCTCCACGAA 2460
 QY 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGlnProValAlaVal 807
 Db 2461 GCAGTTTGGAGAGTACTCAGGAGAGTCTTCAGGTTCCTCCGACGAGCAAGTTGAGTA 2520
 QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
 Db 2521 GCACAGCCCAAGACTACCCAGCCGACCACTTGGCTCTCTCTGTCGACAGTCCACATTCA 2580
 QY 828 AspValAlaSerGlyMetSerSerSerGlyValGlnValValProSerSerSerGlyValHis 847
 Db 2581 GATGTGCTTCAGGATAGATGATGCAATGCAACGTCCTCCACTCCAGTGAAGGCACT 2640
 QY 848 GluGlyArgThrThrLysArgHisTyrArgLysSerValArgSerArgSerArgHisGln 867
 Db 2641 GAAGGAAAGAACTACAAACGCAATTACCAAAATCTGTAAGAGTCGCTTCGACATGAA 2700
 QY 868 LysThrSerArgProLysLeuArgIleLeuValSerAsnLysGlyAspArgValVal 887
 Db 2701 AAAACTCAGCCCAAAATTAAAGATTGTAATGTTTCAMATTAAAGAGACCGAGTAGTA 2760
 QY 888 GluGlyGlnLeuGlnThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly 907
 Db 2761 GAATGTCAATTAGACACTCATATATAGAAATGTTTACATTCMAATTGACCTAGATGCT 2820
 QY 908 AspAsnProGluGlnIleAlaThrIleMetValAsnAsnAspPheIleLeuAlaIleGln 927
 Db 2821 GACAACCCCGAGAGATAGCAACATTTATGGAAACATGACTTTATCTTGCAATATGAG 2880
 QY 928 ArgGlnSerPheValAspGlnValArgGlnIleIleGlnLysAlaAspGlnMetLeuSer 947
 Db 2881 AGAGAGTGTGTTGTGATCAAGTGCAGAAATATTATGAAAAAGCTGATGAATGTGCT 2940
 QY 948 GluAspValSerValGlnProGluGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
 Db 2941 GAGAGTGCAGTGTGAGAACAGAGGGTATCAGGGATGGAGAGCTCTCAAGGAAAGGAT 3000
 QY 968 AspTyrGlyPheSerGlySerGlnLysLeuGlnGlyGlnPheLysGlnProIleProAla 987
 Db 3001 GACTATGGCTTTTCAGGTCTTCAGAAATTTGAGAGAGAGTTCMAACAAACCAATTCCTGCG 3060
 QY 988 SerSerMetProGlnGlnIleGlyIleProThrSerSerLeuThrGlnValValHisSer 1007
 Db 3061 TCTTCCATGCCACAGCAAAATAGGCAATTCCTACAGTTCTTTACTCAAGTTGTTCACTCT 3120
 QY 1008 AlaGlyArgArgPheIleValSerProValProGlnSerArgLeuAlaGlnSerLysVal 1027
 Db 3121 GCGGGAAAGCGGTTTATATGATGTCCTGTGCCAGAAAGCCGATTTACGAAATCAAAAGTT 3180
 QY 1028 PheProSerGlnIleThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
 Db 3181 TTCCCAAGTAAATACAGATACAGTTCCTCTCTACAGCTCAGAGCCCTGGAAATGAC 3240
 QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnAlaPheSerGlnLeuArgArgAla 1067
 Db 3241 TTGTCTCACTCTGCATCATCCCTAGTCTACACAGGCGCTTTTCGAACCTTAAGCGTCC 3300
 QY 1068 GlnMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
 Db 3301 CAAATGACAGAGAGCCCAAAACAGACCTTCAAACTTTAGTCTATACAGGCAACCACTTT 3360
 QY 1088 ProValValProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaThr 1107
 Db 3361 CCAATACCTCTCTTTCTTAAGTACATTCCTGAGTCCCAACACAGCAGCAGCCACA 3420
 QY 1108 AlaProValProAlaThrSerSerProProAsnAspIleSerThrSerValIleGlnSer 1127
 Db 3421 GCACACATCCCTGCACAAAGAGCGCTCTAATGACATTTCCACATCAGTATATCACTCT 3480
 QY 1128 GluValThrValProThrGlnGlnGlyIleAlaGlyValAlaThrSerThrGlyValVal 1147
 Db 3481 GAGGTTACAGTCCCACTGAAGAGGGGATTTGCTGGAATGTCACACAGCAGAGGTGTGTA 3540

QY 1148 ThrSerGlyGlyLeuProIleProProValSerGlnSerProValLeuSerSerValVal 1167
 Db 3541 ACTTCAGGTGTCTCCCACTACACCTGTGTGATACCAAGTACTTCCAGCTAGTT 3600
 QY 1168 SerSerIleThrIleProAlaValValSerIleSerThrThrSerProSerLeuGlnVal 1187
 Db 3601 TCAGTATCACAAATCTCGCACTTCTTCATATATCTACATACCCCGTCACTTCAAGTC 3660
 QY 1188 ProThrSerThrSerGlnIleValValSerSerThrAlaLeuThrProSerValThrVal 1207
 Db 3661 CCCACATCCACATCTGAGATGTTGTTCTAGTACAGCACTATACCTTCACTACAGTT 3720
 QY 1208 SerAlaThrSerAlaSerAlaGlyGlySerThrAlaThrProGlyProLysProProAla 1227
 Db 3721 TCAGCACTTCAGCTCTGACGGGGGAGTCTGCTACCCCAAGGCTCAAGCTCAGCT 3780
 QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValAlaThrLeuThrSerValSer 1247
 Db 3781 GTAGTATCTCAGCAGCAGCAGCAGCAGCTACTGTGGAGGCCACTTAACATCAATTTCT 3840
 QY 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerThr 1267
 Db 3841 ACCACACTTCATTCCTCAAGCAGCAGCTTCACAGCTGCTCATTCAGCTTAGCAGAGTACT 3900
 QY 1268 SerThrProThrLeuAlaGluThrValValValSerAlaHisSerLeuAspLysThrSer 1287
 Db 3901 TCATCTCTACTTATGAGTGAACCGGTGTAGTTAGCCGACACTCATGATTAAGACATCT 3960
 QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
 Db 3961 CATAGCACTAACAATGATTTGGCTTTCTCTCTCTGACCAATCTTCTCTCTCTCTCT 4020
 QY 1308 GlyAlaGlyValSerSerTyrIleSerGlnProGlyGlyLeuHisProLeuValIlePro 1327
 Db 4021 GGACGAGAGAGTGTAGTTATATTTTCAGCGCTGGTGGCTGCATCTTGTGCTATCCA 4080
 QY 1328 SerValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
 Db 4081 TCAGTATAGCTTCTACTCTCATTTCTTCCCAACAGAGAGACCTTCTTCAACCTTTA 4140
 QY 1348 LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal 1367
 Db 4141 TTACCCCAAGTACTAGTATGCCACCTTGGTACAGCTGTGCCATATGTCTCTGCTTA 4200
 QY 1368 GlnGlnThrLeuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
 Db 4201 CAGCAGACACTAATTCATAGTCAAGCTCAACCACTTGTCTTCCCAACCAAGCCCATACT 4260
 QY 1388 HisGlyProGluValAspSerAspThrGlnProLysAlaProGlyIleAspAspIleLys 1407
 Db 4261 CATTTGCTTGAAGTATCTGATATCAACCAACCTCTCTGGAATTTGATGAATGAAG 4320
 QY 1408 ThrLeuGlnGlnLysLeuArgSerLeuPheSerGlnHisSerSerSerGlyAlaGlnHis 1427
 Db 4321 ACTCTAAGAAAAAGCTGGCGGTCTCTGTTCAGTAAACACACTATCTGGAGCTCAGAT 4380
 QY 1428 AlaSerValSerLeuGlnThrThrSerLeuValIleGlnSerThrValThrProGlyIlePro 1447
 Db 4381 GCCCTGTCTCAGCTGAGAGACTCACTAGTATAGAGACTGTACACAGCAGGATCCCA 4440
 QY 1448 ThrThrAlaValAlaProSerLysLeuLeuThrSerThrThrSerThrCysLeuProPro 1467
 Db 4441 ACTATGCTGTGACCAAGCAAACTCTGATCTTCAACCAATTAATCTGTTTACACCA 4500
 QY 1468 ThrAsnLeuProLeuGlnThrValAlaLeuProValThrProValValThrProGlyGln 1487
 Db 4501 ACCAATTTACACATAGAAACAGTTGCTTTCAGCTTACACAGTGTGACACACTGGGCA 4560
 QY 1488 ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys 1507
 Db 4561 GTTCTTACCCCACTGACACTACTACATCAGAGGTGAACCTGGAACTGCTCCTCCAG 4620

Db	1621	ACAGAAATCAACAGACGCTCCCTTCTGCTAGACACCGGACCTACTACTGCTTCTACCACTTCA	1680
QY	621	AlaSerValSerThrGlnValGlnProGlnGlnProGlnAlaaspGlnHisGlnGlnLeu	640
Db	1681	GCTTCAGTTCTTACACAAAGTGAACCTGAAAGAACCTTAGGAGAGATCAACATCAACAACCTA	1740
QY	641	GlnTyrGlnGlnProSerLysSerValLeuSeraspGlyThrValaLapSerGlyGlnGly	660
Db	1741	CAGTACACGACCAACCGATATATCTGTGTATCTGATGAGACGGTTGACAGTGGTCAAGGGA	1800
QY	661	SerSerValPheThrGlnLysSerArgValSerSerGlnGlnThrValSerTyrGlySerGln	680
Db	1801	TCCTTCGTCTTCAACAGAACTCCGAGTAGAGAGCCAAACAGACAGTTATATGTTCCCA	1860
QY	681	HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla	700
Db	1861	CATGAACAGGACCACTTCTACAGGACAGTCCCAAGGACATATACCTTCTACTGTCCAAAGCA	1920
QY	701	GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnGlnSerGlnGly	720
Db	1921	CAGTCTCAGCCCACTGGGGTATATCCACCCCAAGTGGCAGCGGAGAGCAAGGAGT	1980
QY	721	GlnProSerSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740
Db	1981	CAGCCATCTCTCAAGTAGCTTAAACAGGGGTTTCATCTTCCCAACCAATACAAATCCTCAG	2040
QY	741	GlnGlnGlnGlnGlyIleGlnGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer	760
Db	2041	CAGCAGCAGGGAATACACGACAGACAGCCCTCCTCAACAGACAGTCAAGTATTCCTTCA	2100
QY	761	GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	2101	CAGACATAACTCCAGTAGAGGCCACTATCTCACAGCCAGTGAAGTCAAGCTTCAGATCCA	2160
QY	781	GlnValLeuProGlnValSerAlaGlyLysGln-----	791
Db	2161	CAGTCTTGCTCTCAAGTATCAAGCTGGAAAAACGCTTCAGATTTCCAGCCAGTACCACT	2220
QY	791	-----	791
Db	2221	ATCCAAAGCGAACCCTCAATCCAGTTGCGAACACAAACCCTGGTTGTTCCAGTCCACTCT	2280
QY	791	-----	791
Db	2281	GGTGTCTATTCTTCCAGTGGGACAGCGGCTCCTACTCCCTTGCTCCTCAGTACCCT	2340
QY	791	-----	791
Db	2341	GTTCTCTAGATTCCCAATATCAATCTCCTATGTGTACGGCTCAGACAGGTTTCTCATCC	2400
QY	791	-----	791
Db	2401	CTTCCCATCAAAATGGCAGCTGGCATTAATCAAGCTCTGCTCAGCTTGCTTCAATCTGCT	2460
QY	791	-----	791
Db	2461	ACAAACAGCTCGATCCCGGGGATATCAACTGTGTCTTAGTCAAGCTTCCAAACCTTCTG	2520
QY	791	-----	791
Db	2521	CAGCTGTGACTCAGCTGCGAAAGTCAAGTTCAACCCACAGCTCTTACAAACACAGATTCAG	2580
QY	791	-----	791
Db	2581	TCCATGGGAATACACAGTAACTTGGAACAAGCTGTGAGGTTCCACTTCTCTGGAGAT	2640
QY	791	-----	791
Db	2641	GTTCTGTACAGAGGCTTCCACCTGCACTGCAACACAGTACCAAGATTCAAATATTT	2700
QY	791	-----	791
Db	2701	GCTCCTCTTCCACAGTGGCTTCTGTGTTGATCCATCTTACAGTCCATACCCCTCCCATG	2760

QY	791	-----	791
Db	2761	CCGACAGAAAGTACTGGCTACACCTGGGTACTTTCCACAGTGGTGACGCTTATGTGGAA	2820
QY	791	-----	791
Db	2821	TCMAATCTTTAGTTCTCTATAGGTGGTGTCTAGAGAGACAGGTTCAAGTGTCCACCGCA	2880
QY	792	-----	792
Db	2881	GGGAGTTTAGACAGAGCCCGACACTACATCTCCACAGACAGAGTTTGGAGAGTACTCAG	2940
QY	795	GIYVAlserGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln	814
Db	2941	GGAGCTCTCAGGTGTGCTCTCGACAGACCAGTGGCCATACACAGCCCGCAAGCTACCCAG	3000
QY	815	ProThrThrLeuAlaSerSerValAlaSerSerAlaHisSerAspValAlaSerGlyMetSer	834
Db	3001	CCGACACTTTGGGCTCTCTCTGTAGACAGGCAATTCAAGTGTGCTTCAAGTATAGT	3060
QY	835	AspGlyAsnGluAsnValProSerSerSerGlyArgHisGlyArgThrThrIysArg	854
Db	3061	GATGCGCAATGAAAGCGTCCCATCTTCCAGTGAAGCATGAAGAGAAAGACTACAAAACGG	3120
QY	855	HisTyrArgIysSerValArgSerArgSerArgHisGlyIysThrSerArgProIysLeu	874
Db	3121	CATTACCGAAATCTGTATGAGAGTGGCTCTCGACATCAAAAACCTTACCGCCCAAAATT	3180
QY	875	ArgIleLeuAsnValSerAsnIysGlyAspArgValAlaGluCysGlnLeuGluThrHis	894
Db	3181	AGAAATTTGAATGTTTCAATTAATAAGAGACCGAGTAGAGATGTCATTTAGAGACTCAT	3240
QY	895	AsnArgIysMetValThrPheIysPheAspLeuAspGlyAspAsnProGluGluIleAla	914
Db	3241	AATAGGAATAATGGTTACATTCAAAATTGACCTAGTGTGACCAACCCCGAGGAGATGACA	3300
QY	915	ThrIleMetValAsnAsnAspPheIleuAlaIleGluArgGluSerPheValAspGln	934
Db	3301	ACAAATTAACGTAACAATGACCTTATATTCTACCAATAGAGAGAGTGGTGTGTGATCAAA	3360
QY	935	ValArgGluIleIleGluIysAlaAspGluMetLeuSerGluAspValSerValGluPro	954
Db	3361	GTGCAGAGAAATTAATGAAAAAGCTGATGAATGCTCAGTGGAGATGTCAGTGTGAACCA	3420
QY	955	GluGlyAspGlnGlyLeuGluSerLeuGlnGlyIysAspAspTyrGlyPheSerGlySer	974
Db	3421	CAGGGGTGATCGAGGATTTGAGAGGTCTACAGAAAGAGATGCTATGGCTTTACAGTTCT	3480
QY	975	GlnIysLeuGluGlyGluPheIysGlnProIleProAlaSerSerMetProGlnGlnIle	994
Db	3481	CAGAAATTGGAAAGAGAGTCAAAACAACAATTCCTGGCTTCTCATGCCACACAATA	3540
QY	995	GlyIleProThrSerSerLeuThrGlnAlaValHisSerAlaGlyArgArgPheIleVal	1014
Db	3541	GGCAATTCCTACCAAGTTCTTTAACTCAAGTTGTATCTTGGGGAGAGCGGTTATATGTG	3600
QY	1015	SerProValProGluSerArgLeuArgGluSerIysValPheProSerGluIleThrAsp	1034
Db	3601	AGCTCTGTGCCAGAAAGCCGATTACAGAAATCAAAAGCTTTCCCAAGGAATAACAGAT	3660
QY	1035	ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer	1054
Db	3661	ACAGTTGCTGCTCTTACAGCTCAGAGCCCTGGAAAGAACTTGTCATCTGCATCATATCC	3720
QY	1055	LeuSerLeuGlnGlnAlaPheSerGluLeuArgAlaGlnMetThrGluGlyProAsn	1074
Db	3721	CTTAGCTTACCAACAGGCGCTTTTCTAACTTACACGTGCCAAATATACGAAGAGACCTCA	3780
QY	1075	ThrAlaProProAsnPheSerHisThrGlyProThrPheProValAlaProProPheLeu	1094
Db	3781	ACGACACTCCAAACTTTTAGTATACAGACCAACATTTCCAGTAGTATCTCTTTCTTA	3840

Db	6001	AGCGACAAATGAGTCAGATATCCGAAGATGAAGACTTAAAGTTAAGGCTGGACAGACATCAACA	6060
QY	1835	AspLysHisLeuLysGluIleGlnAspLeuGlnSerArgGlnLysHisGluIleGluSer	1854
Db	6061	GATAAACATCTCCAAAGAGATTTCAGAACCTCCAGAGTGCAGGAAGCATGAAATTGAATCT	6120
QY	1855	LeuThrThrlsLeuLysLysValIleProProAlaValIleIleProProAlaIleProLeu	1874
Db	6121	TTGTATATCCAAACTGGGCGAAGGTGCCCCCTGCTGTTATTATTTCCCGAGTGTCCCCCTT	6180
QY	1875	SerGlyArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer	1894
Db	6181	TCAAGGAGAGACACACACACCCTTAAAGCAAGAGCGACGAATCTAGTCCAGACAGATCC	6240
QY	1895	LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaIleSerVal	1914
Db	6241	TTGGGGAATATAAGCCCCCAGCTTTCAGGTAACTCTGTCGTGTCAGATGCAGCTTCAGTC	6300
QY	1915	LeuHisProGlnGlnThrLeuHisLeuProProGlyAsnIleProGluSerGlyGlnIleGln	1934
Db	6301	TTGCACCCCCACGACAGACCCCTCCACCCTCGGCAACATCCACAGATCCGGGCGAAACAG	6360
QY	1935	LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuThrSerAlaPheThrSer	1954
Db	6361	CTGTTACAGCCCCCTTAAAGCATCTCCCTCCAGTACAACTCTATTAGCTTCACCAAGT	6420
QY	1955	AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn	1974
Db	6421	GATGGTGCATTTCAGTATCCAAAGCTTTCTGCTCCAGGTCAAGGAACAGAGACACAAC	6480
QY	1975	ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnIleGlnProProAlaMetThrSer	1994
Db	6481	ACTGTTGGGGCAACAGTGAACAGCCAGCCGCCCAAGTCCAGCTCTGCATGACATCC	6540
QY	1995	SerArgLysGlyThrPheThrAspAspLeuHisIleLysLeuValAspAsnThrAlaArgAsp	2014
Db	6541	AGCAGGAAGGGGACATTCACAGATGACTTGACAAAGTTGGTGGACAAATTGGGCCCCGAGAT	6600
QY	2015	AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrGluIleProGly	2034
Db	6601	GGCAATGATCTCTCAGCGACGAGGAGGAAGCAAGGAGCATGAATTATGAAGGGCCCTGGA	6660
QY	2035	MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly	2054
Db	6661	ATGGCAGAGGAAGTCTCTGCACCTGGGGCACTGTGCATCTCCATCGACCTCGAACCCTGGGT	6720
QY	2055	GlySerAlaProIleSerAlaIleSerAlaThrSerLeuGlyHisIlePheThrLysSerMet	2074
Db	6721	GGCTCTGCCCCCATCTTCGACAGCATCACTCTCTTAGGTCACTTCACCAAGTCAATG	6780
QY	2075	CysProProGlnGlnTyrGlyPheProAlaThrProPheGlyAlaGlnTyrSerGlyThr	2094
Db	6781	TGCCCCCCACAGCACTATGAGCTTTCACACTGCCATTGGGGCTCAATGAGGGAGCG	6840
QY	2095	GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln	2114
Db	6841	GGTGGCCGACGACCAAGCACTGGGCCAGTTCACCACTGTGGGAACCTGCTCTTGGAG	6900
QY	2115	AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg	2134
Db	6901	AATTTTCAAATCAGCAATTTCGAGAAATTCATCAAGCAACCCCCAGGCTCCAACTCGCG	6960
QY	2135	ThrThr 2136	
Db	6961	ACCACT 6966	
RESULT 13			
ACH03794			
ID ACH03794 standard; DNA; 6147 BP.			
XX ACH03794;			
DT 26-SEP-2003 (first entry)			

XX	DNA encoding novel human protein kinase #31.
XX	Human; kinase; ds; gene; cosmetic application; nutraceutical application.
XX	Homo sapiens.
OS	US6541252-B1.
XX	01-APR-2003.
PD	14-MAY-2001; 2001US-00854856.
XX	19-MAY-2000; 2000US-0206015P.
PR	(LEXI-) LEXICON GENETICS INC.
PA	Walke DW, Hilburn E, Donoho G, Turner CA;
XX	MPL; 2003-575927/54.
DR	F-PsDB; ABO44411.
XX	New nucleic acid encoding novel human proteins, useful in cosmetic and
PT	nutraceutical applications.
PS	Disclosure; Page; 11pp; English.
XX	The invention relates to a new isolated nucleic acid encoding a novel
CC	human protein kinase. The nucleic acid is useful in cosmetic and
CC	nutraceutical applications. The present sequence represents DNA encoding
CC	a novel human protein kinase. Note: The sequence data for this patent did
CC	not from part of the printed specification but was obtained in electronic
CC	format directly from USFTO at
CC	segdata.uspto.gov/sequence.html?DocID=6541252B1
SQ	Sequence 6147 BP; 1727 A; 1630 C; 1393 G; 1395 T; 0 U; 2 Other;
Alignment Scores:	
Pred. No.:	0 Length: 6147
Score:	10344.00 Matches: 2048
Percent Similarity:	98.65% Conservative: 0
Best Local Similarity:	98.65% Mismatches: 0
Query Match:	95.67% Indels: 28
DB:	Gaps: 1
US-10-010-720-14 (1-2136) x ACH03794 (1-6147)	
QY	61 MetAspLysASPSeSerArgGlyAlaAlaAlaThrThrThrThrThrGluHisArgPhe 80
Db	1 ATGGACAGGAAGACGCCGTGGGGCGGCCGGACCATCTACCACCACTGAGAGCAGCGTCTTC 60
QY	81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100
Db	61 CGCCGAGGCATCATCTGCACATCCAAATGCACATGCAGACTTCGCGCTTCTCTT 120
QY	101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
Db	121 TCCTCTCCCGACGCCAGCATCCCCGGGCGTGTCGCCAAGAgtCTCACCGAGCCCCAC 180
QY	121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
Db	181 CGGGANAGAACCTGACCGCCACCGCACTTCCAGAGTAGCCAGCAGCTCCAGCGCT 240
QY	141 AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db	241 GCCGCCCTTGAGGAACAGGCCCTGCGGGCCCTGCCCCCTCGACTTCCCAGCAgTACC 300
QY	161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGluProProPro 180
Db	301 AGCAAAGACCGCCCAATGTCCAGCCTTGTGGGAGCAAGAGAGCGCGCGCG 360
QY	181 AlaArgSerGlySerGlyGlySerAlaLysGluProGlnGlnGluArgSerGlnGln 200

Db	361	GCAGAGAGTGGCAGCGCGCGGACGCCACAGGACCCACAGAGAGACGAGCCACGAG	420
Qy	201	GLNAspAsp111eGIuGIuLeuGIuThrLYsAlaValGlyMeSerAsnAspGlyArpHe	220
Db	421	CAGAGATATATCCAGAGGCTGGAGACCAAGCCGTGGCAATGCTACAGATGCCGCTTT	480
Qy	221	LeuLYsPheAsp111eGIuGIuArgGlySerPheLYsThrValTYrLYsGlyLeuAsp	240
Db	481	CTCAAGTTACATCCGAATCGGACAGAGGCTCTTTAAGACGGTCTACAAAGGCTCGAGC	540
Qy	241	ThrGIuThrThrValGIuValAlaTYrCYsGIuLeuGlnAspArgLYsLeuThrLYsSer	260
Db	541	ACTGAACACACCGTGGAGAGTCGCTGGTGAACCTCGAGATCGAAAAATTACAAAGTCT	600
Qy	261	GIuArgGlnArpPheLYsGIuGIuAlaGIuMetLeuLYsGlyLeuGlnHisAspAsn11e	280
Db	601	GAAAGGCACAGATTTTAAAGAGAGAGCTGAATGTTTAAAGGCTTCACAGATCCCAATAT	660
Qy	281	ValArgPheTYrAspSerTrpGIuSerThrValLYsGlyLYsLYsCys111eValLeuVal	300
Db	661	GTTAGATTTTATGATTTCTGGGAATCCACAGTAAAGGAGAAAGAGCATTTGTTGGTG	720
Qy	301	ThrGIuLeuMetThrSerGIuThrLeuLYsThrTYrLeuLYsArgPheLYsValMetLYs	320
Db	721	ACTGAACCTTATAGCGCTCGGAACACTTAAAGCTATCGAAAGGTTTAAAGTGATGAAG	780
Qy	321	11eLYsValLeuArpSerTrpCYsArgGln111eLeuLYsGlyLeuGlnPheLeuHisThr	340
Db	781	ATCAAGATCTTAAGAGAGCTGGTGCCTCAGATCTTAAAGGCTTCACAGTTCTTCATACT	840
Qy	341	ArgTrpProPro111eHisArgAspLeuLYsCYsAspAsn111ePhe111eThrGlyPro	360
Db	841	CGAATCTCCACTATATCTTACCGCATCTTAAATGACAAACATCTTATATCACCGGCCCT	900
Qy	361	ThrGIySerValLYs111eGlyAspLeuGlyLeuAlaThrLeuLYsArgAlaSerPheAla	380
Db	901	ACTGGCTCAGTCAGATTTGAGACCTCGGTCTGGGACCCCTGAAGCGGGCTTTTTCCTC	960
Qy	381	LYsSerVal111eGlyThrProGIuPheMetAlaProGIuMetTYrGIuGIuLYsTYrAsp	400
Db	961	AAGAGTGTGAATGAGTACCCAGAGTTCATGGCCCTGAGATGTATGAGAGAAATATGAT	1020
Qy	401	GIuSerValAspValTYrAlaPheGlyMetCYsMetLeuGIuMetAlaThrSerGIuTYr	420
Db	1021	GAAATCCGTTGACGTTTATGCTTTTGGATGTGATGCTTGAGATGGCTACATCTGAAAT	1080
Qy	421	ProTYrSerGIuCYsGlnAsnAla11eGln111eTYrArgArgValThrSerGIuValLYs	440
Db	1081	CCTTACTCGAGAGTGCACAAATGCTGCACAGATCTACCTCGCGTGCAGCCAGTGGGGTGAAG	1140
Qy	441	ProAlaSerPheAspLYsValAla11eProGIuValLYsGIu111eGlyGlyCys111e	460
Db	1141	CCAGCCAGTTTTCACAAAGTAGCAATCTCTGAAGAGAGAAATTAATTGAAGATGACATA	1200
Qy	461	ArgGlnAsnLYsAspGlnArgTYrSer111eLYsAspLeuLeuAsnHisAlaPhePheGln	480
Db	1201	CGACAAAACAAAGATGAAGATATTCATCAAAAGACCTTTGAAACCAAGCCCTTCTTCCAA	1260
Qy	481	GIuGIuThrArgValArgValGIuLeuAlaGIuGIuAspArgLYsGIuLYs111eAla111e	500
Db	1261	GAGGAAACAGAGAGTACGGGTGAATTCACAGAGAGATGTGAGAGAAAAAATATGCCATA	1320
Qy	501	LYsLeuTYrPheLeuArg111eGlnAsp111eLYsLYsLeuLYsGIuTYrLYsAspAsnGIu	520
Db	1321	AAATTTATGCTACGATTAAGATATTAAAGAAATTAAGGGAAAAATCAAAAGATTAAGAA	1380
Qy	521	Ala111eGIuPheSerPheAspLeuGIuArgAspValProGIuAspValAlaGlnGlyMet	540
Db	1381	GCTATTGAGTTTTCTTTGATTTTGAAGAGAGATGTCACAGAGATGTTTGAACAAGAAATG	1440
Qy	541	ValGIuSerGIuTYrValCYsGIuGIuLYsPhe111eTYrThrMetAlaLYsAla111eLYsAsp	560
Db	1441	GTAAGTCTGGGTATGTCTGTGAAGAGTATACAAAGCCATGGCTTAAAGCTATCAAAAGAC	1500

Qy	561	ArgValSerLeu111eLYsArgLYsArgGlyGIuGlnArgGlnLeuValArgGIuGlnGlnGIu	580
Db	1501	AGAGTATCATTTATTTAAAGAGAAACGAGACGCGGCTTGTTGTTGGAGAGACAAAGA	1560
Qy	581	LYsLYsLYsGIuGlnGIuSerSerLeuLYsGlnGlnValGIuGlnSerSerAlaSerGln	600
Db	1561	AAAAAAAAGCAGAAAGAGAGAGCTCTCAAAACAGCAGGTAAACAATCAGTGTCTCCAG	1620
Qy	601	ThrGIy111eLYsGlnLeuProSerAlaSerThrGly111eProThrAlaSerThrThSer	620
Db	1621	ACGGAATCAAGACGCTCTCTTGTGCTAGACACCGGACATACCTACTGCTTACCACTTCA	1680
Qy	621	AlaSerValSerThrGlnValGIuProGIuGIuProGIuAlaAspGlnHisGlnGlnLeu	640
Db	1681	GCTTCAGTTTCTACACAGATGAACCTGAGAACCTTAGAGCAGATCAACATCAACAATA	1740
Qy	641	GlnTYrGlnGlnProSer111eSerValLeuSerAspGlyThrValAspSerGIyGlnGly	660
Db	1741	CAGTACAGCAACCCAGATATCTGTGTATCTGATGAGAGGTTGACAGTGGTCAAGGA	1800
Qy	661	SerSerValPheThrGIuSerArgValSerSerGlnGlnThrValSerTYrGlySerGln	680
Db	1801	TCTCTGTCTTCACAGAAATCTCGAGTACGCCAACAGACAGATTATATGTTATCTCCAA	1860
Qy	681	HisGlnGlnAlaHisSerThrGIuThrValProGIuHis111eProSerThrValGlnAla	700
Db	1861	CATGAACAGGACATTTCTACAGGACAGCTCCAGGAGCATATCTTCACTGTCTCAAGCA	1920
Qy	701	GIuSerGlnProHisGlyValTYrProProSerSerValAlaGlnGlyGlnSerGlnGly	720
Db	1921	CAGTCTCAGCCCATGAGGATATCCACCTCAAGTGTG-----	1959
Qy	721	GlnProSerSerSerSerLeuThrGIuValSerSerSerGlnPro111eGlnHisProGln	740
Db	1959	-----	1959
Qy	741	GlnGlnGlnGlnGly111eGlnGlnThrAlaProProGlnGlnThrValGlnTYrSerLeuSer	760
Db	1960	---CAGCAGGGAATACAGCAGACAGCCCTCTCAACAGACATGCAATTCATCTTCA	2016
Qy	761	GlnThrSerThrSerSerGIuAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	2017	CAGCATCAACCTCCAGTGAAGCCACTACTGACACAGCAGTGAAGTCAAGCTCCCA	2076
Qy	781	GlnValLeuProGlnValSerAlaGlyLYsGlnSerThrGlnGlyValaSerGlnValAla	800
Db	2077	CAAGTCTTGCTCAAGTATCAGTGTGAAGAAACAGAGTACTCAGAGAGTCTCAGTGTGCT	2136
Qy	801	ProAlaGlnProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer	820
Db	2137	CTGCAAGAGCACTTCGACAGTACAGACGCCCAAGCTAACCCAGCCACACTTGGCTTCC	2196
Qy	821	SerValAspSerAlaHisSer	

QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgIuIleIleGlu 940
DB 2497 GACTTATCTTAGCAATAGAGAGAGAGTCCGTTTGATCCAGGCAAAATTATGAA 2556
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGluAspGlnGluLeu 960
DB 2557 AAAGCTGATGAAAGTCACTGAGAGATGATGATGAAACAGAGGGGTGATCAGGGAGTGG 2616
QY 961 GluSerLeuGlnGluLysAspAspTyrGluPheSerGlySerGlnLysLeuGluGlu 980
DB 2617 GAGAGCTTACAGAGAAAGAGATGACTATGCTTTTCAAGTTCTCAGAAATTGGAAAGAGAG 2676
QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
DB 2677 TTCAACACACCAATCTCGGCTCTTCATGCAACAGCAAAATAGGCATTTCTACAGATTCT 2736
QY 1001 LeuThrGlnValValHisSerAlaGlyValArgArgPheIleValSerProValProGluSer 1020
DB 2737 TTAACTCAAGTTGTTCACTTCGCGGGAAGCGGTTTATGAGATCCTGTCGCAAGAGC 2796
QY 1021 ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
DB 2797 CGATTAGAGAGATCAAAAGTTTTCCTCAAGTGAATTAACAGATACAGTTGCTGCTCTACA 2856
QY 1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnAla 1060
DB 2857 GCTCAGAGCCCTGGAATGAAGTCTGCTCCTCATCATCATCTTATGCTTACACAGAGCC 2916
QY 1061 PheSerGluLeuValArgValAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe 1080
DB 2917 TTTTCTGAACTTACAGCGTCCCAATGACAGAAAGAGCCCAVAGACACCTCCCAAACTTT 2976
QY 1081 SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGlyVal 1100
DB 2977 AGTATACAGAGCAACATTTCCAGTAGTACCTCTCTTTCTTATGATGATTCCTGAGATC 3036
QY 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
DB 3037 CCAACCCAGACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 3096
QY 1121 SerThrSerValIleGlnSerGlnValThrValProThrGluGluGlyTLeAlaGlyVal 1140
DB 3097 TCCACATCAGTAAATTCAGTCTGAGGTTACAGTGCCTGACGAAAGGGGATTCGAGATT 3156
QY 1141 AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer 1160
DB 3157 GCCACCCAGACAGAGGTGAGTCACTCAGTGTCTCCCATACCACTGTGTCTGATCA 3216
QY 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
DB 3217 CCAAGTACTTCCAGGCTGATTCAAGTATCAAAATCACTGCACTGCTCAATATCTACT 3276
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla 1200
DB 3277 ACATCCCGGTCACTTCAAGTCCCAATCAGTCACTGATCGTGTCTTCAAGTACACCA 3336
QY 1201 LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
DB 3337 CTGATTCCTCAGTACAGTTCAGCAACTTCAGCTTCGACGGGGGAGTACTGCTACC 3396
QY 1221 ProGlyProLysProProAlaValValSerGlnAlaAlaGlySerThrThrValGly 1240
DB 3397 CCAAGTCTTAGGCTCCAGGCTGATCTATCTCAGGAGGACGACGACGACGACGACGACG 3456
QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
DB 3457 GCCACATTAACATCAGTTTCTACCACTTCATTCCTCCAAAGCAGCAGCTTCACAGCTGCC 3516
QY 1261 IleGlnLeuSerSerSerThrSerThrProThrLeuAlaGluThrValValValSerAla 1280
DB 3517 ATTACGCTTACGAGGATCTTCTACTCTTACTTTAGCTGAACCGTGTGATTTAGCGCA 3576
QY 1281 HisSerLeuAspLysThrSerHisSerThrThrGlyLeuAlaPheSerLeuSerAla 1300

DB 3577 CACTCACTAGATAGACATCTTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 3636
QY 1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerTyrIleSerGlnProGlyGly 1320
DB 3637 CCATCTTCTCTCTCTCTCTCTGAGCAGAGAGTCTTATATTTCTTCTGACCGCTGGGG 3696
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
DB 3697 CTGCATCTTGTGCTATTCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 3756
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
DB 3757 GGAAGCTTACTTACACTTATTAATCCCAAGTACTAGTATCCCACTTGTGATCAGAGCT 3816
QY 1361 ValAlaAsnValProAlaValGlnIleThrLeuIleHisSerGlnProGlnProAlaLeu 1380
DB 3817 GTTGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3876
QY 1381 LeuProAsnGlnProHisThrHisCysProGluValAspSerAspThrGlnProLysAla 1400
DB 3877 CTTCACCAAGCAGCCCACTACTCATTTGCTGAAGTATGATTCATACACCAACCAAGCT 3936
QY 1401 ProGlyTLeuAspIleIleThrLeuGluGluLysLeuArgSerLeuPheSerGluHis 1420
DB 3937 CCTGGAATTGATCACAATAAGACTTGAAGAAAGCTGGGCTCTGTTCAAGTAAAC 3996
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrSerLeuValIleGluSer 1440
DB 3997 AGCTCATCTGAGCTCAGCATGCTCTGCTCTACTGAGAGCTCATCATGCTATGAGAGC 4056
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
DB 4057 ACTGTACACAGGCAATCCCAACTACTGCTGTGCAACCAAGCAAACTCTGACTTACAC 4116
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
DB 4117 ACAAGTACTGCTTACACCAACCAATTTACCATAGCAACAGTGTCTTGCCAGTTACA 4176
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
DB 4177 CCAAGTGTACACTGGGAGAGTTTCTTACCCAGTCAAGCTACTATCATCAGAGGTGAA 4236
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
DB 4237 CTTGAACTGCTCTCTCAAGCCACTTACTTAAAGCTCGGAGTCCGACGAGTGTACT 4296
QY 1521 GluLeuProAlaGlyThrLeuProSerGluGlnLeuProProPheProGlyProSerLeu 1540
DB 4297 GAAGTCCAGAGGATCTTACCCAGGAGAGCTGCCACTTTTCCAGGACCTTCTCTA 4356
QY 1541 ThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeuArgThrLeuSerPro 1560
DB 4357 ACCAGTCCAGCAACCTCTAGAGATCTTATCTCATTTGAGAAAGACCTTAGTCCA 4416
QY 1561 Glu**IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
DB 4417 GAGATKATCACAGTACTTTCGCGTGTGCTGTGTCATCAGGGCGCTCCAAACAGCATC 4476
QY 1581 ThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGluGlyProValLeu 1600
DB 4477 ACAGAGCAGAGAACACACCTCAAGAAAGGTCTTCTCAAGTCAAGAAAGGCGCTGTCTTA 4536
QY 1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
DB 4537 GCACTAGTTCAGAGCTGCTGCTTTTATGATGGAGCATTTCAAGGTTCTGTTGCACCA 4596
QY 1621 AspGlyAlaGlnLysGluGlyLysAsnLysSerGluAspAlaLysSerValHisPheGlu 1640
DB 4597 GACGGTCCCAAGAAAGGTGAAATTAAGTACAGAGTGAAGATCTGTTCAATTTTAA 4656
QY 1641 SerSerThrSerGluSerSerValLeuSerSerSerProGluSerThrLeuValLys 1660

Db 4657 TCACGACCTCAGAGTCTCAGTGTATCATAGTACAGAGTACCTTGGTAA 4716
 QY 1661 ProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGluSerAlaHis 1680
 Db 4717 CCAGAGCCGAAATGCAATACATCCCTGATCTCTCAGATGTGCAGAGTGTCCAC 4776
 QY 1681 LysThrThrAlaSerGluAlaLysSerAspThrGlyGlnProThrLysValGlyThrPhe 1700
 Db 4777 AAAATCTAGCTCTCAGAGGCAAGTCAACACTGGGAGCCTTACCAAGTTGAGCTTTT 4836
 QY 1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGluAspLys 1720
 Db 4837 CAGGTGACAACTACAGCAAAAGAGTGGTCTCTTCTGTATCAAAAATCAGAGCAAG 4896
 QY 1721 IleThrAspThrLysLysGluGlyProValAlaSerProProPheMetAspLeuGln 1740
 Db 4897 ATCATGTACACAAAGAAAGAGGACCAAGTGCATCTCTCTTATGATTTGAGACAA 4956
 QY 1741 AlaValLeuProAlaValIleProLysLysGlyLysProGluLeuSerGluProSerHis 1760
 Db 4957 GCTGTTCTCTCTGCTGTATACCAAGAAAGAGAGCTGAACTGTACAGAGCTTCAAT 5016
 QY 1761 LeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspValAspAspGly 1780
 Db 5017 CTAAATGGGCGGTCTTGACCCGAGGCGCTTTTAAAGTAAAGATGTGATGATGAT 5076
 QY 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGluProSerHis 1800
 Db 5077 TCCGGTATCTCACACTGCCCATCAGTGAAGTAAAGAGCTTCTTCAAGAGATCTA 5136
 QY 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSerSerAspAsnGlySerAsp 1820
 Db 5137 AGTCAAAGCCTTAGTAATTCATTTAACTCTTACATGAGTGAAGCAATGAGTCAT 5196
 QY 1821 IleGluAspGluAspLeuLeuGlnLeuArgArgLeuAlaAspLysHisLeuLysGlu 1840
 Db 5197 ATCAAGATGAGACTTAAAGTTAGAGCTGCAGACATCAAGATTAACATCTCAAGAG 5256
 QY 1841 IleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLeuLysThrLysLeuGly 1860
 Db 5257 ATTCAGAGCCTCGAGAGTCCGCGCAGAGATCAAAATTGAAATTTTGTATACCAAACTGGG 5316
 QY 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgArg 1880
 Db 5317 AAGGTGCCCTGCTGTATTATTCGCCCAAGCTCTCCCTTTCAGGAGAGAGAGAGAG 5376
 QY 1881 ProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerPro 1900
 Db 5377 CCCACTAAAGCAAAAGGAGCAAAATCTAGTCGAAGCAGTTCTTGGGGAATMAAGCCCC 5436
 QY 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
 Db 5437 CAGCTTTCAGGTACCTGTGTGTGTCAGAGTCAAGCTTCAAGCTTTCAGCCCCAGAGACC 5496
 QY 1921 LeuHisProProGlyAsnIleProGluSerGlyGlnAsnGlnLeuGlnProLeuLys 1940
 Db 5497 CTCACACCTCTGTCAGACATCCAGAGTCCGGGAGCAATCAGCTTACAGCCCCCTTAAG 5556
 QY 1941 ProSerProSerSerAspAsnLeuLysSerAlaPheThrSerAspGlyAlaIleSerVal 1960
 Db 5557 CCATCTCTCTCCATGACAACTCTATTTCAGCTTCAACGATGATGTGTGCATTTCAGTA 5616
 QY 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
 Db 5617 CCAGAGCTTCTGTGTGTCAGGTCAAGAGAACAGCAAGCAACTGTTGGGGCAAGAGTG 5676
 QY 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
 Db 5677 AACAGCCCAAGCCCCCAAGCTCAGCTCCGTGCCATGACGTCCAGAGAGAGGAGCACTTC 5736
 QY 2001 ThrAspAspLeuHisLysLeuValAspAsnThrAlaArgAspAlaMetAsnLeuSerGly 2020
 Db 5737 ACAGATGACTTGCAAGATTGGTAGACAAATTGGGCGCCGAGATGCATGAAATCTCTCAGGC 5796

QY 2021 ArgArgGlySerLysGlyHisMetAsnTyrGlnGlyProGlyMetAlaArgLysPheSer 2040
 Db 5797 AGGAGAGAGAGAGCAAGAGGACATGAAATTATGAGGGCCCTCGAATGGCAAGAAAGTTCTT 5856
 QY 2041 AlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGlyLysSerAlaProIleSer 2060
 Db 5857 GCACCTGGGCAACTGTGCAATCTCCATGACCTCGAACTGGGGGTCTGCCCCCAATCTCT 5916
 QY 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnTyr 2080
 Db 5917 GCAGCATCAGCTACCTCTCTAGTCACTTCAACAGATCATGTGCCCCCAGACAGATGT 5976
 QY 2081 GlyPheProAlaThrProPheGlyAlaGlnTyrSerGlyThrGlyProAlaProGln 2100
 Db 5977 GCGTTTCCAGCTACCCCAATTGGCGCTCAATGAGTGGGACGGGTGGCCAGCACACAG 6036
 QY 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
 Db 6037 CCATTGGCCAGTTCCAACTGTGGAACTGCTCTTGCAGAATTTCAACATTCAGCAAT 6096
 QY 2121 LeuGlnLysSerLysSerAsnProProGlySerAsnLeuArgThrThr 2136
 Db 6097 TTGCAGAAATCATCATCAGCAACCCCGAGGCTCCAACTGGGGACCACT 6144

RESULT 14
 ACH03791
 ID ACH03791 standard; DNA; 6426 BP.
 XX
 AC ACH03791;
 XX
 DT 26-SEP-2003 (first entry)
 XX
 DE DNA encoding novel human protein kinase #28.
 XX
 KW Human; kinase; ds; gene; cosmetic application; nutraceutical application.
 XX
 OS Homo sapiens.
 XX
 EN US6541252-B1.
 XX
 PD 01-APR-2003.
 XX
 PF 14-MAY-2001; 2001US-00854856.
 XX
 PR 19-MAY-2000; 2000US-0206015P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hilbun E, Donoho G, Turner CA;
 XX
 DR WPI; 2003-575927/54.
 DR P-PSDB; ABO44408.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutraceutical applications.
 PS
 PS Disclosure; Page: 11pp; English.
 XX
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents DNA encoding
 CC a novel human protein kinase. Note: The sequence data for this patent did
 CC not from part of the printed specification but was obtained in electronic
 CC format directly from USPO at
 CC seqdata.uspto.gov/sequence.html?DocID=6541252B1
 XX
 SQ Sequence 6426 BP; 1789 A; 1714 C; 1455 G; 1465 T; 0 U; 3 Other;

Alignment Scores:
 Pred. No.: 0 Length: 6426
 Score: 10287.50 Matches: 2048
 Percent Similarity: 94.42% Conservative: 0

Best Local Similarity: 94.42%
 Query Match: 95.15%
 DB: 8
 Mismatches: 0
 Indels: 121
 Gaps: 2

US-10-010-720-14 (1-2136) x ACH03791 (1-6426)

QY 61 MetAspIysAspSerArgIylAlaAlaAlaThrThrThrThrThrGluHisArgPhePhe 80
 DB 1 ATGGACAGAGACAGACCGCTGGGGCGCGCGACCACTACCACTAGACACCGCTTTC 60
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 DB 61 CGCGGAGCGCTCATCTCGACTCCCAATGCACATGCACTGAGAGCTTCCCGGCTTCTT 120
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGlnProHis 120
 DB 121 TCCCTGCCACCGCCAGCATCCCGCGGCTGCTCCGAGAGTGTCTCCACCGAGACCCAC 180
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 DB 181 CGGGAAGAGACCGTGACCGCCACCGCCACTTCCAGGTAGCCGACGAGCCCTCCAGCGCT 240
 QY 141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 DB 241 GCCGCCCTGGGGAACAGCGCGTGCAGCGCCCTGCTCCGACTGTCTCCAGCAAGTACC 300
 QY 161 SerIlyAspArgProValSerGlnProSerLeuValGlySerIlyGluGluProProPro 180
 DB 301 AGCAAAACCGCCAGTGTCCAGCTTACCTTGTGGGAGCAAGAGAGAGCCCGCCGCG 360
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaGlyGluProGlnGluGluArgSerGlnGln 200
 DB 361 GCGGAAGTGGCAACGCGCGCGCGCGCAAGCCCAAGAGACCAAGAGAGAGAGAGAGAG 420
 QY 201 GlnAspAspIleGluGluLeuGluThrIlyAlaValGlyMetSerAsnAspGlyArgPhe 220
 DB 421 CAGATGATATCGAAGAGCTGAGACCAAGCGCGTGGAAATGTCTTAACGATGCGCGCTT 480
 QY 221 LeuIlyPheAspIleGluIleGlyArgGlySerPheIlyThrValIlyIlyGlyLeuAsp 240
 DB 481 CTCAAGTTTGACATCGAATCGCAAGAGCTCTTTAAAGACGCTTACAAAGGCTTGAC 540
 QY 241 ThrGluThrThrValGluValAlaIleProCysGluLeuGlnAspArgIlyLeuThrIlySer 260
 DB 541 ACTGAACACACCGTGAAGTGCCTCGTGTGATCGACGAGATCGAAATTTAAACAAGCT 600
 QY 261 GlnArgGlnArgPheIlyGluGluAlaGluMetLeuIlyGlyLeuGlnHisProAsnIle 280
 DB 601 GAGAGGACAGATTAAAGAAAGAGCTGAAATGTTAAAGGCTTCCACATCCCAATATT 660
 QY 281 ValArgPheIlyArgSerIlyGluSerThrValIlyGlyIlyIlyCysIleValLeuVal 300
 DB 661 GTTAATTTTATGATTTCTCGGAATCCACAGTAAAGAAAGAAAGAGATGCTTTTGATG 720
 QY 301 ThrGluLeuMetThrSerGlyThrLeuIlyThrIlyLeuIlySerPheIlyValMetIly 320
 DB 721 ACTGAACCTTAAGACGCTCGAACAACCTTAAGATCTGAAAGATTTAAAGTATGATGAG 780
 QY 321 IleIlyValLeuArgSerIlyCysArgGlnIleLeuIlyGlyLeuGlnPheLeuHisThr 340
 DB 781 ATCAAAAGTTCTAAAGAGCTGGTGGCGGTGAGATCTTAAAGGCTTCCATGCTTCTACT 840
 QY 341 ArgThrProProIleIleHisArgAspLeuIlyCysAspAsnIlePheIleThrGlyPro 360
 DB 841 CGAACTCCACCTATATTCACCGCAATCTTAAATGACACATCTTATATCCCGGCGCT 900
 QY 361 ThrGlySerValIlyIleGlyAspLeuGlyLeuAlaThrLeuIlyArgAlaSerPheAla 380
 DB 901 ACTGCTCATGTCAGATTCAGACCTCGGTCTGGAGAACCTGAAAGGCGCTTCTTTGCC 960
 QY 381 IlySerSerValIleGlyThrProGluPheMetAlaProGluMetIlyGluGluIlyIlyAsp 400
 DB 961 AAGAGTGTATGATGATACCCAGAGTTCATGCGCCCTGAGATGTATGAGAGAAATATGAT 1020

QY 401 GluSerValAspValIlyAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluIly 420
 DB 1021 GAATCCGTGACGTTATATGCTTTGGATGTGATGCTGAGATGCTACATCTGAATAT 1080
 QY 421 ProIlySerGluCysGlnAsnAlaAlaGlnIleIleIlyArgArgValThrSerGlyValIly 440
 DB 1081 CCTTACTCGGAGCGCAAAATGTGCACAGANTCTCCGTGCTGACCAAGTGGGTGAG 1140
 QY 441 ProAlaSerPheAspIlyValAlaIleProGluValIleGluIleIleGluGlyCysIle 460
 DB 1141 CCAGCCAGTTTACAAAGTACGATATTCGAAAGTGAAGAAATTTATGAAGATGCATA 1200
 QY 461 ArgGlnAsnIlyAspGluArgIlySerIleIlyAspLeuLeuHisAlaPhePheGln 480
 DB 1201 CGACAAACCAAGATGAATATTCATCAAGACCTTTGAACCATGCTTCTTCCAA 1260
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspGlyGluIlyIleAlaIle 500
 DB 1261 GAGGAAACAGAGATACCGGTAGAAATTAGCAAGAAAGATATGAGAAAAAATAGCCATA 1320
 QY 501 IlyLeuThrPheLeuArgIleGluAspIleIlyLeuIlyGlyIlyIlyIlyAspAsnGlu 520
 DB 1321 AATTTATGCTTACGATATTAAGATATTAAGAAATTAAGGAAATTAAGAAATATGAA 1380
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 DB 1381 GCTATGATGTTTCTTTGATTTAGAGAGAGATGCTCCAGAGATGTTGCAACAAGAAATG 1440
 QY 541 ValGluSerGlyIlyValIlyCysGluGlyAspHisIlySerPheMetAlaIlyAlaIleAsp 560
 DB 1441 GTAGAGTCTGGATATGCTGTGAAGATGATCACAAAGCCATGCTAAAGCTATTAAGAC 1500
 QY 561 ArgValSerLeuIleIlyAsArgIlyGluGlnArgGlnLeuValArgGluGluGlnGlu 580
 DB 1501 AGATATCATTTAATTAAGAGAAACGAGAGAGCGGAGTGTGATCGGAGAGACAGAA 1560
 QY 581 IlyIlyIlyGluGluGluSerSerLeuIlyGlnGlnValGluGlnSerAlaSerGln 600
 DB 1561 AAAAAAAGCAGAGAGAGAGAGAGCTCAAAACAGAGTAGAAACATCAAGTCTTCCAG 1620
 QY 601 ThrGlyIleIlyGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 DB 1621 ACAGAAATCAAGAGCTCCCTCTGTGACGACCGGACATCTTACTGCTTACCACTTCA 1680
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
 DB 1681 GCTTCAGTTTCTACAAAGTGAAGACCTGAAGACCTGAGGACGATCAACATCAACACTA 1740
 QY 641 GlnIlyGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 DB 1741 CAGTACGACCAACCAATATCTGTGTTATCTGATGGAGCGGTTGACAGTGGTCAAGGA 1800
 QY 661 SerSerValPheThrGluSerArgValSerGlnGlnThrValSerIlyGlySerGln 680
 DB 1801 TCCCTGCTTCAAGAAATCTGAGTGAAGCAACCAAGATTTCTATAGTGTCCCA 1860
 QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 DB 1861 CATGAACGACATTTCTACAGGACAGATCCAGGCAATATACCTTACTACTGCCAACA 1920
 QY 701 GlnSerGlnProHisGlyValIlyProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 DB 1921 CAGTCTCAGCCCATGGGTATATCCACCTTCMACTG----- 1959
 QY 721 GlnProSerSerSerLeuThrGlyAlaSerSerSerGlnProIleGlnHisProGln 740
 DB 1959 ----- 1959
 QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnIlySerLeuSer 760
 DB 1960 ---CAGCAGGAAATACAGACAGACGCGCTCTCTCAACAGACAGTCAAGTATCACTTCA 2016

QY 761 GlnThrSerThrsSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2017 CAGACATCAACCTCCAGTAGAGGCACTACTGCACAGCCAGTAGTACGCTCAAGTCTCA 2076
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 2077 CAAGTCTGCTCAAGTATCACTGGAAGAAACAGGGCTTCCACCTGCAGCTGCACACAG 2136
QY 791 ----- 791
Db 2137 TACCCAGGAGATTCAATATGTGCTCCCTCTCCAACTGCGCTTCTGTTCATTCATTC 2196
QY 791 ----- 791
Db 2197 ACAAGTCTAACCCCTCCCATGCGCAGACAGAGTACTGCGTACCTGGGACTTTCACCA 2256
QY 791 ----- 791
Db 2257 GTGGTGCAGCCTTATGTGATCAATCAATCTTTAGTTCCTATGGGTGTGTAGAGAGACAG 2316
QY 791 ----- 791
Db 2317 GTTCAAGTGTCCAGCCAGAGAGAGGAGTTAGACACAGCCCACTACATCTCCACAGCA 2376
QY 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGluProValAlaVal 807
Db 2377 GCAGTCTTTCAGAGACTCAGGAGAGTCTCTCAGGTGTCTCTGCAGAGCCAGTGTGACATA 2436
QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2437 GCACAGCCCAAGCTACCCAGCCAGCCAGCTTGGCTTCTCTGTAGACAGTGCACATTCA 2496
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyAlaArgHis 847
Db 2497 GATGTGCTTCAAGTATAGATAGATGATGCGCATAGAACGCTCCACTTCCAGTGAAGGAT 2556
QY 848 GlnGluValThrThrLysArgHisLysArgLysSerValArgSerArgSerArgHisGlu 867
Db 2557 GAAGAGAAAGACTACAAACCGGCAATTACCGAAATCTGTAGAGAGTCGCTCGACATGAA 2616
QY 868 LysThrSerArgProLysLeuArgLysLeuAsnValSerAsnLysGlyAspArgValVal 887
Db 2617 AAAAATTCCAGCCCAAAATTTAGAAATTTGAAATGTTCTCAATTAAGAGACCGGATGTA 2676
QY 888 GluCysGlnLeuGlnThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly 907
Db 2677 GAATGTCAATTAGAGACTCAATTAAGAAATGTTTACATTCAATTTGACCTTAGATGT 2736
QY 908 AspAsnProGlnGluLysLeuAlaThrLysMetValAsnAsnAspPheLeuAlaAlaGlu 927
Db 2737 GACAAACCCCGAGAGATAGCAACATATATGTGAAACATGACTTATTTAGCAATAGAG 2796
QY 928 ArgGluSerPheValAspGlnValArgGluLeuLeuLysAlaAspGluMetLeuSer 947
Db 2797 AGAGAGTCGTTTGTGATCAAGTGCAGAAATTAATTAAGAAAGCTGATGAAGTCTCAGT 2856
QY 948 GluAspValSerValGluProGlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
Db 2857 GAGGATGTCAAGTGTGAGAACAGAGGGGTGATCAGGAGATTGAGAGACTTACAAAGAAAGAT 2916
QY 968 AspTyrGlyPheSerGlySerGlnLysLeuGlnGlyLysLysGlnProIleProAla 987
Db 2917 GACTATAGCTTTTCAGGTTCTCAGAAATTTGAAAGAGAGTTCAAACCAACATTCCTCGCG 2976
QY 988 SerSerMetProGlnGlnIleGlyLysLeuProThrSerSerLeuThrGlnValAlaHisSer 1007
Db 2977 TCTTTCATGCAACAGCAATAGGATTCCTTCAACAGTCTTAACTCAAGTGTTCATTC 3036
QY 1008 AlaGlyArgArgPheLeuValSerProValProGlnSerArgLeuArgGlnSerLysVal 1027
Db 3037 GCGGGAAGGGGTTTATAGTGAATCTCGTCCAGAAAGCCGATTAAGAGATTAATAAGTT 3096
QY 1028 PheProSerGluLysThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047

Db 3097 TTCCCAAGTAAATACAGATACATGCTGCTCTACAGCTCAGGCGCTCGAATGAAC 3156
QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAla 1067
Db 3157 TTGTCTCACTCGATCATCTCTTATGTCTACACAGGCTTTTCTGAATTTAGAGCTGCG 3216
QY 1068 GlnMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
Db 3217 CAATGACAGAGAGGCCAAACAAACAGACCTCCAACTTATGATCAATACAGACCAACATT 3276
QY 1088 ProValValProProPheLeuSerSerLysLeuAlaGlyValProThrThrAlaAlaThr 1107
Db 3277 CCAATGATCCTCTTCTTATAGTACATGCTGTGAGTCCCAACACAGCAGCAGCAGCA 3336
QY 1108 AlaProValProAlaThrSerSerProProAsnAspLysSerThrSerValIleGlnSer 1127
Db 3337 GCACAGTCCCTGCAAGACAGAGAGCCCTCTTAATGACATTTCCACATCAGTAAATTCAGTCT 3396
QY 1128 GluValThrValProThrGlnGlnGlyLysLeuAlaAlaThrSerThrGlyValVal 1147
Db 3397 GAGTTACAGTGCACCTGAAAGAGGGATGCTGAGATTGCCACACAGCAGGTGTGTA 3456
QY 1148 ThrSerGlyLysLeuProIleProProValSerGlnSerProValLeuSerSerValVal 1167
Db 3457 ACTTCAGGTGTCTCCCATACACACTGTGTGTGAATCACAGTATCTTCCAGCGTAGTT 3516
QY 1168 SerSerIleThrIleProAlaValValSerLysSerThrThrSerProSerLeuGlnVal 1187
Db 3517 TCAAGTATCACAATACCGGAGTGTGTCTCAATATCTACTAATCCCGTCACTTCAGTCT 3576
QY 1188 ProThrSerThrSerGlnLysValValSerSerThrAlaLeuLysProSerValThrVal 1207
Db 3577 CCCACATCCACATGTGAGTGTGTTGTTCTAGTACAGACAGTATCTTCACTGTAACAGTT 3636
QY 1208 SerAlaThrSerAlaSerAlaGlyLysSerThrAlaThrProGlnProLysProProAla 1227
Db 3637 TCAGCACTTCAGCTCTGCAAGGGGAGATGATGCTACCCCAAGGTCTTAACCCCTCCAGCT 3696
QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
Db 3697 GTAGTATCTCAGCAGCAGCAGGAGCAGCAGTACTGAGGAGCAGATTAATCATCGTTCT 3756
QY 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerLysLeuLeuSerSerThr 1267
Db 3757 ACCACCACTTCATTTCCACACACAGCCTTCAAGCTGTCCATTCAGCTTACAGCAGTACT 3816
QY 1268 SerThrProThrLeuAlaGluThrValValValSerAlaHisSerLeuAspLysThrSer 1287
Db 3817 TCTACTCTCACTTATAGCTGAACCGGTGTAGTGAAGCCACATCACTAGATTAAGACATCT 3876
QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
Db 3877 CATAGCAGTACCACTGGATTGGCTTCTCTCTCTGCACATCTTCTCTCTCTCTCTCT 3936
QY 1308 GlyAlaGlyValSerSerTyrLysSerGlnProGlyGlyLeuHisProLeuValIlePro 1327
Db 3937 GGAGCAGAGATGTCTAGTATATTTCTCAGCTGAGGGGCTGATCTTTGGTCAATTC 3996
QY 1328 SerValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
Db 3997 TCAGTATAGCTTCTACTCTCTATTTCTTCCCAAGCAGCAGAGCTTACTTACACCTTTA 4056
QY 1348 LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal 1367
Db 4057 TTACCCCAAGTACTTATGATCCACCTTGTGACAGCTGTGTCATGCTGTCTGTGA 4116
QY 1368 GlnGlnThrLeuLysHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
Db 4117 CAGCAGACCTATTTATGATGAGCTCAACAGCTTCTCTCCCAACAGCAGCCCACTACT 4176
QY 1388 HisCysProGlnValAspSerAspThrGlnProLysValProGlyLysAspAlaIleLys 1407

Db	4177	CATTGTCCTGAAGTAGATTCTGATACACAAACCCAAAGCTCTGGAAATTGATGACATAAG	4236
QY	1408	ThrlengluGluIysLeuArgSerLeuPheSerGluHisSerSerSerGlyAlaGlnHis	1427
Db	4237	ACTTAGAAGAAAAGCTGGCGGTCTCTTTCAGTGAACACAGCTCATCTGGAGCTCAGCAT	4296
QY	1428	AlaSerValSerLeuGluThrSerLeuValIleGluSerThrValThrProGlyIlePro	1447
Db	4297	GCCTGTGCTCACGTGAGAGACTCACTGATCAGAGAGCACTGTCAACACGAGCATGCCA	4356
QY	1448	ThrThrAlaValAlaProSerLysLeuLeuThrSerThrThrSerThrCysLeuProPro	1467
Db	4357	ACTACTGCTGTGGACCAAGCAAACTCTTCACTTACCAAAAGTACTTCTTACCAACA	4416
QY	1468	ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValValThrProGlyGln	1487
Db	4417	ACCAATTTCACACTAGGAACAGTTGCTTTGGCAATTACACAGTGGTCAACACTGGGCCA	4476
QY	1488	ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys	1507
Db	4477	GTTTCTACCCAGTCAGACACTACTACATCAGAGTGAACCTGGAACTGCTCCCTCCAG	4536
QY	1508	ProProLeuThrLysAlaProValLeuProValGlyThrGluLeuProAlaGlyThrLeu	1527
Db	4537	CCACTCTTAACCTAAGGCTCCGAGTCTGCATGTGGTACTGAATCTCAGAGGTACTCTA	4596
QY	1528	ProSerGluGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnGlnProLeu	1547
Db	4597	CCGACGAGCACTGCCACTTTTCCAGAGCTTCTTAACCCAGTCCGACCAACTCTA	4656
QY	1548	GluAspLeuAspAlaGlnLeuArgArgThrLeuSerProGlu**IleThrValThrSer	1567
Db	4657	GAGGATCTTGTATGCTCAATTGAGAAGAACCTTAGTCCAGAGATKATCACAGTCACTTC	4716
QY	1568	AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGluAlaGlyThrGlnPro	1587
Db	4717	GCGGTGTGCTCGTGTCCATGCGCGCTCCACACACAAACACAGAGAGCAACACACTT	4776
QY	1588	GlnLysGlyValSerGlnValIleGluGlyProValLeuAlaThrSerSerGlyValGly	1607
Db	4777	CAGAAAGGATGTTTCTCAAGTCAAGAGAGGCCCTGTCTTAGAACCTAGTCAGAGAGCTGT	4836
QY	1608	ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyAlaGlnLysGluGly	1627
Db	4837	GTTTTTAAGATGGAGCATTTTCAGGTTTCTGTGTGACAGACAGCGGCCCAAGAAAGAGGT	4896
QY	1628	LysAsnLysSerGluAspAlaLysSerValHisPheGluSerSerThrSerGluSerSer	1647
Db	4897	AAAAATTAAGTCAGAAAGATGCAAAAGCTCTTCAATTGATCCACACACTCGAAGTCTCTCA	4956
QY	1648	ValLeuSerSerSerSerProGluSerThrLeuValLysProGluProAsnGlyIleThr	1667
Db	4957	GGCTATACAGTAGTAGTCCAGAGGTACCTTGTGTGAACCAAGACCGAATGGCATTAAC	5016
QY	1668	IleProGlyIleSerSerAspValProGluSerAlaHisLysThrThrAlaSerGluAla	1687
Db	5017	ATCCCTGGATATCTTTCAGATGTGCCAGAGAGTCCCAAAACTACTGCTCAGAGGCA	5076
QY	1688	LysSerAspThrGlyGlnProThrLysValGlyArgPheGlnValThrThrThrAlaAsn	1707
Db	5077	AAGTCAGACACGCGGAGCTTACCAAGTTGGACCTTTTCAGGTATCAACTACGCAAC	5136
QY	1708	LysValGlyArgPheSerValSerLysThrGluAspLysIleThrAspThrLysLeuGln	1727
Db	5137	AAAGTGCGTCGTTTCTCTGTATCAAAAACCTGAGGACAGATCACTGACACAAAGAAAGA	5196
QY	1728	GlyProValAlaSerProPheMetAspLeuGlnGlnAlaValLeuProAlaValIle	1747
Db	5197	GGACCAAGTCATCTCTCTTTATATGATTTGAACAAGCTGTTCTTCGTGCTGATA	5256
QY	1748	ProLysLysGluLysProGluLeuSerGluProSerHisLeuAsnGlyProSerSerAsp	1767
Db	5257	CCAAAGAAAGAGAGAGCTGAACGTCAAGGCTTCAACATCTAAATGAGGCGGCTTCTTCAAC	5316

QY	1768	ProGlu1aAla1aPheLeuSerAaArgAspValaAspAspGlySerGlySerProHisSerPro	1787
Db	5317	CCGAGAGCCCTTTTAAAGTACGAGATGAGGAAAGATGGTTCGGTGGATCCACCTGGCC	5376
QY	1788	HisGlnLeuSerSerLysserLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer	1807
Db	5377	CATAGCTGAGCTCAAAAGACCCTTCCTAGCCAAATTAAGTCAAAAGCCTTAGTAATTCA	5436
QY	1808	PheAsnSerSerTyrmcSerSerSerAspAsnGlnSerAspIleGlnAspGlnAspLeuLys	1827
Db	5437	TTTAACCTCCTTTCATCAGATGACGACAAATGAGTCAATATCGAATGAAGACTTAAAG	5496
QY	1828	LeuGlnLeuAaArgAaArgLeuAaArgAspLysHisLeuLysGlnIleGlnAspLeuGlnSerAaG	1847
Db	5497	TTAGAGCTGGCAGCAGCTACGAGATTAACATCTCAAAAGAGATTGAGACCTGGCAGAGCTGGC	5556
QY	1848	GlnLysHisGlnIleGlnSerLeuTyrrThyThyLeuGlnLysValProProAlaValIle	1867
Db	5557	CAGAAAGCATGAAATGTAATCTTGTAATACCAAACTGGGCAAGGGGCCCTGCTGTATAT	5616
QY	1868	IleProProAlaAlaProLeuSerSerGlyAaArgAaArgAaProThrLysSerLysGlySer	1887
Db	5617	ATTCCCCAGCTGCTCCCTCCCTTTCAGGAGAAAGACGACGCCACTAAACCAAGAGAGC	5676
QY	1888	LysSerSerAaArgSerSerSerLeuGlyAAsnLysSerProGlnLeuSerGlyAsnLeuSer	1907
Db	5677	AAATCTAGTCAGAGCAGTCTCTGGGGAATTAAGCCGCCAGCTTTCAGGTAACCTGTCT	5736
QY	1908	GlyGlnSerAlaAlaIleSerValLeuHisProGlnGlnThrLeuHisProProGlyAsnIle	1927
Db	5737	GGTCAGAGTGAAGCTTCAGTCTTCCACCCCGACGACCCCTCCACTCTGGGAAATC	5796
QY	1928	ProGlnSerGlyGlnAsnGlnLeuLeuGlnProLeuLysProSerProSerSerAspAsn	1947
Db	5797	CCAGAGTCGGGACGAATCAGCTGTTACAGCCCTTAGCCATCTCCCTCAGAGACAC	5856
QY	1948	LeuTyrrSerAlaPheThrSerSerAspGlyAlaAlleSerValProSerLeuSerAlaProGly	1967
Db	5857	CTCTATTGAGCTTCACAGTGAATGGTGCATTTCACTACCAACCTTTCTGCTCCAGGT	5916
QY	1968	GlnGlyThrSerSerThrAsnThrValGlyAlaThrValAsnSerGlnAlaGlnAla	1987
Db	5917	CAAGGACCCAGAGACCAAAACACTGTGGGGCAACAGTGAACACCCAGCCGCCAAAGCT	5976
QY	1988	GlnProProAlaMetThrSerSerAaArgLysGlyThrPheThrAspAspLeuHisLysLeu	2007
Db	5977	CAGCTCTCGCATGACGTCCAGAGGAAAGGGACATTCACAGATGACTTGGACAAAGTTG	6036
QY	2008	ValaAspAsnTyrAlaArgAspAlaMetAsnLeuSerGlyAaArgGlySerLysGlyHis	2027
Db	6037	GTAAGCAATTGGGCCGAGATGCCATGATCTTCAGGACAGAGAGGAAAGGAGGAC	6096
QY	2028	MetAsnTyrrGlnGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCysIle	2047
Db	6097	ATGAATTTATGAGGGCCCTGGAATGGCAAGAAATCTCTGGACCTGGGCAACTGTGATC	6156
QY	2048	SerMetThrSerAsnLeuGlyGlySerAlaProIleSerAlaAlaSerAlaThrSerLeu	2067
Db	6157	TCCATGACCTCAACCTGGGTGGCTCTGCCCCCATCTTGCAGCATCAGCTACTCTCTA	6216
QY	2068	GlyHisPheThrLysSerMetCysProProGlnGlnTyrrGlyPheProAlaThrProPhe	2087
Db	6217	GGTCACTTCAACCAAGCTATATGCCCCCACAAGCAGTANGGCTTCCAGTACCCCAATT	6276
QY	2088	GlyAlaGlnTyrSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlnPheGlnPro	2107
Db	6277	GGGAGCTCATGAGAGTGGAGCGGTGGCCCAAGACCAAGCCACTTGGCAAGTCCAACT	6336
QY	2108	ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsn	2127
Db	6337	GTGGGAATCGCTCTTGGAGAAATTTCAACATCAGCAATTTTGGACAAATTCATCAGCAAC	6396

QY 2128 ProProGlySerAsnLeuArgThrThr 2136
 Db 6397 CCCCAGGCTCCAACTGCGGACCACT 6423
 RESULT 15
 ACH03788
 ID ACH03788 standard, DNA, 6885 BP.
 AC ACH03788;
 XX
 AC ACH03788;
 XX
 DT 26-SEP-2003 (first entry)
 DE DNA encoding novel human protein kinase #25.
 XX
 XX Human; kinase; ds; gene; cosmetic application; nutraceutical application.
 XX
 OS Homo sapiens.
 XX
 PN US6541252-B1.
 XX
 PD 01-APR-2003.
 XX
 PF 14-MAY-2001; 2001US-00854856.
 XX
 PR 19-MAY-2000; 2000US-0206015P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hilbun E, Donoho G, Turner CA;
 XX
 DR WPI: 2003-575927/54.
 DR P-RSDB; ABO44405.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutraceutical applications.
 XX
 PS Disclosure; Page: 11pp; English.
 XX
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents DNA encoding
 CC a novel human protein kinase. Note: The sequence data for this patent did
 CC not from part of the pinned specification but was obtained in electronic
 CC format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=6541252B1
 XX
 SQ Sequence 6885 BP; 1881 A; 1873 C; 1543 G; 1585 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 6885
 Score: 10211.00 Matches: 2048
 Percent Similarity: 88.20% Conservative: 0
 Best Local Similarity: 88.20% Mismatches: 0
 Query Match: 94.44% Indels: 274
 DB: 8 Gaps: 2
 US-10-010-720-14 (1-2136) x ACH03788 (1-6885)
 QY 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrGlnHisArgPhePhe 80
 Db 1 ATGAGCAAGACAGACCCGTGGGCGCGCGGACCACTACCACTGAGCACCCGCTTCTTC 60
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100
 Db 61 CGCCGAGCGCTCATCTGCGACTCCAAATGCCACTGAGAGCTTCCGGGCTTCTCTT 120
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGlnProHis 120
 Db 121 TCCCTGCCCCAGGCCAGCATCCCGCGGCTGTCGCCGAGAGTCTCCACCGAGGCCCAAC 180
 QY 121 ArgGlnGlnThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 181 CGGGAAGACACCTGACCCGCCACCTTCCAGGTAGCCGACGAGAGCTCCAGCCGCT 240

QY 141 AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 241 GCCGCCCTGGGGAGACAGCCGCTCGCGGCGCTCCCTCCAGTGTCCCCAGCAGTACC 300
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGlnProProPro 180
 Db 301 AGCAAAAGACCCGCCAGTGTCCAGCTTGTGGGGAGCAAGAGAGAGCGCGCCG 360
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGlnProGlnGlnGlnLysSerGln 200
 Db 361 GCGAAGAGTGGAGCGCGCGCGCGGAGCCCAAGAGCCACAGAGAGAGACGACGAC 420
 QY 201 GlnAspAspIleGlnGlnLeuGlnThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
 Db 421 CAGGATGATTCGAAGACGTGAGACCAAGCCGCGGGAATGTCTAAAGATGCGCTTT 480
 QY 221 LeuLysPheAspIleGlnIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
 Db 481 CTCAAGTTGACATCGAATCGGCAAGGCTCTTTAAGACGTCTACAAAGTCTGAC 540
 QY 241 ThrGlnThrThrValGlnValAlaTyrCysGlnLeuGlnAspArgLysLeuThrLysSer 260
 Db 541 ACTGAAACCAACCGTGAAGTGCCTGCTGTGAATCGACGATCGAAATTAACAAGCT 600
 QY 261 GlnArgGlnArgPheLysGlnGlnAlaGlnMetLeuLysGlyLeuGlnHisProAsnIle 280
 Db 601 GAGAGCAGAGATTTAAGAAAGAGCTGAATGTTAAAGGCTTCAACATCCCAATATT 660
 QY 281 ValArgPheTyrAspSerTyrPylSerThrValLysGlyLysCysIleValLeuVal 300
 Db 661 GTTAGATTATATGATTCCTGGGAATCCACAGTAAAGAAAGAGTGTATTGTTG 720
 QY 301 ThrGlnLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
 Db 721 ACTGAACCTTAGACGTCTGGAACACTTAAACGATCTGAAGAGTTTAAAGTATGAT 780
 QY 321 IleLysValLeuAspSerTyrPylSerThrValLysGlyLysGlnPheLeuHisThr 340
 Db 781 ATCAAACTTCAAGAACTGTGCTGCAAGTCTTAAAGTCTTCAAGTCTTCAACATCT 840
 QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 Db 841 CGAATCCACCTTCAATTCACCGGAGCTTAATGTACAAACATCTTATCACCGGCTT 900
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
 Db 901 ACTGGCTCACTCAAGATTGAGACCTCGTCTGGCAACCTGAAGCGGCTTCTTTGCC 960
 QY 381 LysSerValIleGlyThrProGlnPheMetAlaProGlnMetTyrGlnGlnLysTyrAsp 400
 Db 961 AAGAGTGTATAGTACCCCAAGTTCATAGCCCTGAGATGATAGGAGAAATATAGAT 1020
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlyTyr 420
 Db 1021 GAATCGTTGACGTTTATGCTTTTGGGATGTGATGCTTATGATGGTATCATCTGAATAT 1080
 QY 421 ProTyrSerGlnCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
 Db 1081 CCTTACTCGGAGTGCACAAATGCTGCACAGATCTACCGTGCCTGACCACTGGGGTGAAG 1140
 QY 441 ProAlaSerPheAspLysValAlaIleProGlnValLysGlnIleIleGlnGlyCysIle 460
 Db 1141 CAGGCCAGTGTGACAAAGTAGCAATTCGAAAGTAGAAGAAATTTTGAAGATGCATCA 1200
 QY 461 ArgGlnAsnLysAspGlnArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
 Db 1201 CAGCAAAACAAAGATGAAGATATTCATCAAGACCTTTTGAACCATGCTTCTTCCAA 1260
 QY 481 GlnGlnThrGlyValArgValGlnLeuAlaGlnGlnAspGlyGlnLysIleAlaIle 500
 Db 1261 GAGGAACAGAGGTAGCGGTATAGATTAAGCAAGAGATGATGAGAAATAATAGCATAT 1320

QY 501 LysLeuTrpLeuArg11eglnuAsp11elysLysLeuLysGlyLysTrpLysAspAsnGlu 520
 DB 1321 AAMTATNGCTAGTATGTAAGATATTAAGAAAATTAAGGAAATATCAAAAGATATGA 1380
 QY 521 Ala11eglupheserPheaspleuGluArgAspValProGluAspVal11aglnGluMet 540
 DB 1381 GCTATTCAGTTCTTCTTGTGATTTAGAGAGATGTCCTCAAGAGATGTTGACAAAGAAATG 1440
 QY 541 ValGluSerGlyTrpValCysGluGlyAspHisLysThrMetAlaLysAla11elysAsp 560
 DB 1441 GTAGAGCTGGGTATGCTGTGAAGGTGATCAACAGCATGGCTTAAGCTATCAAGAC 1500
 QY 561 ArgValSerLeuLeuLysArgLysArgGluGlnArgGlnLeuValArgGluGlnGlu 580
 DB 1501 AGAGTATCATTAATTAAGAGAAACGAGACAGCGGAGTTGGTACGGAGAGAAAGAA 1560
 QY 581 LysLysLysGlnGlnGlnSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln 600
 DB 1561 AAAAAAAGCAGAGAGAGACAGCTCAAAACAGCAGTAGAACAATCCAGTGTCCAG 1620
 QY 601 ThrGly11elysGlnLeuProSerAlaSerThrGly11leProThrAlaSerThrThrSer 620
 DB 1621 ACGAGATCAAGACAGCTCCCTCTGTAGACACGGCATACCTACTGCTTACCACTCA 1680
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGlu11aAspGlnHisGlnGlnLeu 640
 DB 1681 GCTTCACTTCTACACAGTATGAACCTGAGAACTTAGGAGCATCAACATCAACAACTA 1740
 QY 641 GlnTrpGlnGlnProSer11eserValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 DB 1741 CAGTACAGCAACCCAGTATATCTGTGTATCTGATGGAGCGGTTGACAGTGTCAAGG 1800
 QY 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTrpGlySerGln 680
 DB 1801 TCCCTGTCTTCAACAGATCTCGAGTAGAGCCAGCAACAGACATTTCAATAGTGTCCAA 1860
 QY 681 HisGlnGlnAlaHisSerThrGly11ThrValProGlyHis11leProSerThrValGlnAla 700
 DB 1861 CATAAACAGGCACATTTCTACAGGCACAGCTCCAGGCAATATACCTTCTACTGTCCAGCA 1920
 QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 DB 1921 CAGTCTCAGCCCATGGGGTATATCCACCTCAAGTGTG----- 1959
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnPro11eglnHisProGln 740
 DB 1959 ----- 1959
 QY 741 GlnGlnGlnGly11eglnGlnThrAlaProProGlnGlnThrValGlnTrpSerLeuSer 760
 DB 1960 ---CAGCAGGGAATACAGCAGACAGACCCCTCTCTCAACAGACAGTGCATTTACTTCA 2016
 QY 761 GlnThrSerTrpSerSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 DB 2017 CAGACATCAACCTCCAGTAGAGGCCACTACTGCACAGCAGTAGAGTCAACCTCAAGTCCA 2076
 QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
 DB 2077 CAAGCTTGCCTCAAGTATCAGCTGGAACAGCTTCCAGTTCCAGCCAGTCAAACT 2136
 QY 791 ----- 791
 DB 2137 ATCCAGGCGAACCCTCAGATCCCAAGTTCGACACAAACCTCGGTTGTTCCAGTCCACT 2196
 QY 791 ----- 791
 DB 2197 GATGCTCATTTCTTCCAGTGGGACAGCGCTCCTACTCCCTTGTCTCAGTACCT 2256
 QY 791 ----- 791
 DB 2257 GTCTCTCAGATTCCCATATCAACTCCTCATGTGTCTACGGCTCAGACAGGTTTCTCATCC 2316
 QY 791 ----- 791

DB 2317 CTTCCTCATCAATGCGACGTGGCATTTACTCAGCTTGTGTACGTTGGCTTCACTGTCT 2376
 QY 791 ----- 791
 DB 2377 ACAACAGCTGGCATTCGGGGGGTATCACTGTGGTTCTTAGTACGCTTCAACCTTCTG 2436
 QY 791 ----- 791
 DB 2437 CAGCTGTACGTAGCTGCCAAGTCAAGTTCACGCCACAGCTCTCAACACAGCAGTTCAG 2496
 QY 791 ----- 791
 DB 2497 TCCATGGGAATACCACTTAACCTTGGACCAAGCTGTGAAGTTCACTTCTCTGTGAGAT 2556
 QY 791 ----- 791
 DB 2557 GTTCTGTACAGGGCTTCCCACTCGACTGCCACACAGTACCCAGAGATTCAAATATT 2616
 QY 791 ----- 791
 DB 2617 GCTCCCTCTTCAACGTGGCTTCTGTTTGATCATCTTCAAGTCTTAMCCCTCCATG 2676
 QY 791 ----- 791
 DB 2677 CCGACAGAACTACTGCTACACTGGGTACTTCCCAAGTGTGACGCTTATGTGAA 2736
 QY 791 ----- 791
 DB 2737 TCAATTTTCTTACTTCTTATGGGTGTGTAGAGACAGGTTCAGTGTGCCAGCCAGCA 2796
 QY 792 -----SerThrGln 794
 DB 2797 GGGAGTTTACAGACAGCCCACTACATCTCCACAGACAGATTTTGAGAGTACTCAG 2856
 QY 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
 DB 2857 GGAATCTCTCAGGTGCTCTCTGACAGACAGTGGTACATAGACAGCCCAAGTACCCAG 2916
 QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 DB 2917 CCGACCACTTGTGGCTCTCTGTAGACAGTGCATTCAGATGTGCTTCCAGTATGAGT 2976
 QY 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGlnGlyArgThrThrLysArg 854
 DB 2977 GATGGCATAGGAACGCCATCTTCCAGTGAAGGATGAAGGAAGTAACTAACAACCG 3036
 QY 855 HisTrpArgLysSerValArgSerArgSerArgHisGlnLysThrSerArgProLysLeu 874
 DB 3037 CATTAACGAAATCTGTAGAGAGTGCCTCGACATGAAAAAATTCACGCCCAAAATTA 3096
 QY 875 Arg11leuAsnValSerAsnLysGlyAspArgValAlaGluCysGlnLeuGluThrHis 894
 DB 3097 AGAATTTTGAATGTTTCAATTAATAGAGACCGAGTAGTAGATCATTTAGAACATCAT 3156
 QY 895 AsnArgLysMetVal11ThrPheLysPheAspLeuAspGlyAspAsnProGlu11leuAla 914
 DB 3157 AATAGGAAATAGTTTCAATTAATAGAGACCGAGTAGTAGATCATTTAGAACATCAT 3216
 QY 915 Thr11MetValAsnAsnAspPhe11leuAla11eglnArgGluSerPheValAspGln 934
 DB 3217 ACAATTTATGTGAACAATGACTTTATCTTACCAATAGAGAGAGTGGTTGTGATCAA 3276
 QY 935 ValArgGlu11le11eglnLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
 DB 3277 GTGGCAGAAATTTATTAAGAAAGCTGATGAATTCCTCAGTACGATGTCAAGTGAACCA 3336
 QY 955 GluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAspAspTrpGlyPheSerGlySer 974
 DB 3337 GAGGGTATTCAGGGATTTGAGAGTCTACAAAGAAAGATGATGCTTGTTCAGTGTCT 3396
 QY 975 GlnLysLeuGluGluPheLysGlnPro11leProAlaSerSerMetProGlnGln11le 994

Db 3397 CAGAAATTGAGAGAGAGTTCACCAACCAATTCTGCGCTTCATGCCACAGCAATA 3456
QY 995 GlyLeuProThrSerSerLeuThrGlnValHisSerAlaGlyArgArgPheIleVal 1014
Db 3457 GGCATTCTCCACAGATCTTTAACTCAAGTGTTCATTCGCGGAGAGCGGTTTAATG 3516
QY 1015 SerProValProGluSerArgLeuArgGluSerValPheProSerGluIleThrAsp 1034
Db 3517 AGTCCGTGCGCAGAAAGCCGATTCAGAGATCAAAAGTTTCCACAGTAATAACAT 3576
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSer 1054
Db 3577 AAGTGTGCTGCTTACAGCTCAGAGCCCTGGAAGAACTGTCTCACCTGCAATCC 3636
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgAlaGlnMetThrGluGlyProsn 1074
Db 3637 CTGTGTCACACAGAGCTTTTCTGAACCTTGACCTGCCCAAAATACAGAAAGACCA 3696
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProPheLeu 1094
Db 3697 ACAGACCTCCAACTTAGTCATACAGACCAACATTTCCAGTAGTACCTCTTCTTA 3756
QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 3757 AGTAGCATGCTGAGAGTCCACACACAGACAGACCAAGCAAGCTCCGCAACAGC 3816
QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 3817 AGCCCTCTTATGACATTTCCACATCAGTATTCAGTGTGAGTTACGTGCTCCACGAA 3876
QY 1135 GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLeuProIle 1154
Db 3877 GAGGGAGTGTGTGAGTGGCCACACAGACAGAGTGTGTTACTTCAGGAGGTCTCCCA 3936
QY 1155 ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 3937 CCACCTGTGTGATACACAGACTCTTCCAGCTTCAAGTGTTCACAAATACACATCTGCA 3996
QY 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 3997 GTGTGTCCAAATATCTACTACATCCCGTCACTTCAAGTCCCAATCCACATCTGAAATC 4056
QY 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 4057 GTTGTTCCTAGTACAGACACTGTATCTTCAGTACAGTTTCAGCACTTCAGCTTCGCA 4116
QY 1215 GlyGlySerThrAlaThrProGlyProGlyProProAlaValValSerGlnGlnAla 1234
Db 4117 GGGGGCAGTACTGTACCCAGGCTTAAAGCTCCACCTGATGATCTCAGCGAGCGCA 4176
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 4177 GGCAGCACTACTGTGGAGCAATTAACATCAGTTCTACACCACTTCAATCCCAAGC 4236
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlu 1274
Db 4237 ACAGCTTCACAGCTGCTCATTCAGCTTACAGCAGAGTCTTCACTCTTAAGTGAA 4296
QY 1275 ThrValValValSerAlaHisSerLeuAspLeuThrSerHisSerSerThrThrGlyLeu 1294
Db 4297 ACCGTGTAGTACCGCACACTCACTGATTAAGACATCTATAGCACTAACCTGGAATTG 4356
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
Db 4357 GCTTCTCCCTCTGTGACATCTTCTCTCTCTCTGAGACAGAGGTCTTAGTAT 4416
QY 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4417 ATTTCAGAGCTGTGGGTGACATCTTGTGTCATTCATCAGTATGAGTTCTTACCTCT 4476
QY 1335 IleLeuProGlnAlaAlaGlyProThrThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 4477 ATTCTTCCCAAGCAGACCTACTTCTTACACCTTTATTAACCCAGTACTTACTATC 4536
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4537 CCACCTGTGTACAGCGCTGTGCCAATGTGCTGTGTAAGCAGACACTAATTCATGT 4596
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisGlyProGluValAspSer 1394
Db 4597 CAGCTTCACAGACTTGTCTTCCCAACAGAGCCCACTCATTTGCTCGAAGTAACTTCT 4656
QY 1395 AspThrGlnProValAlaProGlyIleAspAspIleIleThrLeuGluGluIleLeuArg 1414
Db 4657 GATACACACCAAGCTCCTGAAATGATGACATTAAGACCTTGAAGAAAGCTGGCG 4716
QY 1415 SerLeuPheSerGluHisSerSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 4717 TCTGTGTGATGAAACAGCTCATCTGAGACTGAGCATGCTGCTGCTCACTGGAGACC 4776
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 4777 TCACTAGTCATAGAGACACTGTACACACAGCAATCCCACTACTGCTGTGACAGCAAGC 4836
QY 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 4837 AAACCTCGACTTACACAGAGTACTTGTACCAACACATTTACACTAGAACCA 4896
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 4897 GTTGTCTTGCAGTTACACAGAGTGTACACCTGTGGCAAGTTTTCACCCAGTACGACT 4956
QY 1495 ThrThrSerGlyValLysProGlyValThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 4957 ACTACATCAGAGAGGAACCTGGAAGTCTCTCCCAAGCACTCTTAAAGCTGCG 5016
QY 1515 ValLeuProValGlyThrGluLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 5017 GTGTGTCAAGTGGGTACTGAATCTTCAGAGGATCTTACCAAGCGAGAGCTGCCACT 5076
QY 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 5077 TTTCCAGAGCTTCTCTTAAACCAAGTCCCAAGCACTTAAAGGATCTTGTGATGCTCAATTG 5136
QY 1555 ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValGlyProValSerMet 1574
Db 5137 AGAAGAACACTTGTGTCAGAGATKATACAGTATCTTGTGGTGTCTGTGATC 5196
QY 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 5197 GCGGCTCCAAACAGCAATCACAGAAAGCAAGCAAGCTCAGAAAGGTGTTCTCAAGTC 5256
QY 1595 LysGluGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 5257 AAAAGAAAGCCCTCTCTCAGCACTGATTCAGAGCTGTGTATTAAGAGGAGCAATTT 5316
QY 1615 GlnValSerValAlaIleAspGlyAlaGlnLysGlyLysAsnLysSerGluAspAla 1634
Db 5317 CAGGTTTGTGTGACAGACAGCTGCCCAAGAAAGAGGTAATAATAGTCAGAAAGATGCA 5376
QY 1635 LysSerValHisPheGluSerSerThrSerGluSerSerValLeuSerSerSerPro 1654
Db 5377 AAGTGTTCATTTTGAATCCAGCACTCAGAGTCTCTCAGTGTCTATCAAGTATGATGCA 5436
QY 1655 GluSerThrLeuValLysProGluProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 5437 GAGAGTACCTGTGTGAACCAAGCGAATGGATTAACCATCTCTGTATCTTCTAGAT 5496
QY 1675 ValProGluSerAlaHisIleThrThrAlaSerGluAlaLysSerAspThrGlyGlnPro 1694
Db 5497 GTGCCAGAGAGTCCCAAAAATCTACTGCTCAGAGCAAGTCAACACTGGGAGCT 5556
QY 1695 ThrLysValGlyArgPheGlnValThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 5557 ACCAAGGTGAGAGTTTTCAGGTGACAACTACAGCAAAAGAGTGTGTTCTCTGTGA 5616

QY 1715 SerLySThrGlnuAspLysIleThrAspThrIlySLeuGluGlyProValAlaSerProPro 1734
 DB 5617 TCAAAAACAGAGGACCAAGATCACTGACACAAAGAAAGAGGACAGGGGATCTCTCTCT 5676
 QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGluLysProGlu 1754
 DB 5677 TTATGGATTGGACAAGCTGTTCTTCTGCTGATACCAAGAAAGAGGAGGCTGAA 5736
 QY 1755 LeuSerGlnProSerHisLeuAsnGlnProSerSerAspProGlnAlaIlePheLeuSer 1774
 DB 5737 CTGTGAGAGCTTCAACATCTAAATGAGCGCTTCTTGACCCGAGGCGCTTTTAAAGT 5796
 QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
 DB 5797 AGGAGATGTGATATGTTTCCGATGACCACTCCGCCCATGAGCTGAGCTGAAAGAGC 5856
 QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSer 1814
 DB 5857 CTTCCTAGCCAGATCTAAAGTCAAAAGCCTTAGTAATTCATTTAACTCTTACATGAGT 5916
 QY 1815 SerAspAsnGlnSerAspIleGlnuAspGlnuAspLeuLysLeuGlnuAspArgLeuArg 1834
 DB 5917 AGGACACATAGTCAGATATCGAAGATGAAGACTTAAAGTTAGAGTCGACGACGCTAGCA 5976
 QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
 DB 5977 GATTAACATCTCAAAAGATTTCAAGACCTGACAGCTGCCAGAAAGCATGAATTGAATCT 6036
 QY 1855 LeuLysThrLysLeuGlnLysValProProAlaValIleIleProProAlaAlaProLeu 1874
 DB 6037 TTGTATACCAAACTGGGCAAGTGCCTCCCTGCTGTATTTATTCCTCCAGCTGCTCCCTT 6096
 QY 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
 DB 6097 TCAGGAGAAAGACACACACCCACTAAAGCAAAAGCAGCAAAATCTAGTCGAGAGCATTTCC 6156
 QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
 DB 6157 TTGGGGAATTAAGACCCCGAGCTTCAAGTAACTGTCTGTGTCAGAGTGACGCTTCAAGTC 6216
 QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnuSerGlyGlnAsnGln 1934
 DB 6217 TTGCACCCCGACAGACCTTCACCTCTGCAACATCCAGAGTCCGGGAGAGATCAG 6276
 QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
 DB 6277 CTGTTACAGCCCTTAAGCCATCTCCCTCCAGTACAACTCTTAATCAGCTTCAACAGT 6336
 QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
 DB 6337 GATGCTGCATTTCAATACCAAGCTTCTGTCTCCAGSTCAAGAGACGACGACCAAAAC 6396
 QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaIleGlnAlaGlnProProAlaMetThrSer 1994
 DB 6397 ACTGTTGGGGCAACGTAACAGCCCAAGCCCAAGCTCAGCTCTCCCATGACGTCC 6456
 QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnThrAlaArgAsp 2014
 DB 6457 AGCAGAGAGGGACATTCAAGATGACTTGACAAAGTTGGTAGCAATTGGGCGCGAGAT 6516
 QY 2015 AlaMetAsnLeuSerGlyArgArgLysSerLysGlyHisMetAsnLysGlnGlyProGly 2034
 DB 6517 GCCATGATCTCTCAGGACGAGAGAGGAGCAAAAGGCAATGAATTAGAGGCGCTGGA 6576
 QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuLysLysMetThrSerAsnLeuGly 2054
 DB 6577 ATGCAAGAGAGTCTCTGACACTGGCAACTGTCATCTCATGACCTCCGAACCTGGGT 6636
 QY 2055 GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlnHisPheThrLysSerMet 2074
 DB 6637 GGCTCTGCCCATCTCTGACACATCAGTACTCTTAGAGTCACTTCAACAAAGTCTAGG 6696
 QY 2075 CysProProGlnGlnIlyGlyPheProAlaThrProPheGlyAlaGlnIlyPserGlyThr 2094

DB 6697 TGCCCCCAGACAGATAGCTTTCCAGCTACCCCATTTGGCGCTCAATGAGTGGAGC 6756
 QY 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
 DB 6757 GGTGGCCCAACACACAGCCACTTGCCCAAGTTCAAACCTGTGGAACTGCTCTTGACG 6816
 QY 2115 AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg 2134
 DB 6817 AATTCAATCAGCAATTGGAGAAATCCATCAGCAACCCCAAGCTCCAACTGCGG 6876
 QY 2135 ThrThr 2136
 DB 6877 ACCACT 6882

Search completed: September 25, 2004, 05:23:57
 Job time : 1804 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 25, 2004, 04:08:33 ; Search time 7148 Seconds
(without alignments)
8923.562 Million cell updates/sec

Title: US-10-010-720-14
Perfect score: 10812
Sequence: 1 MSGAAEKQSSRPSGLFLSP.....NISLQKSIISNPPGSNLRRT 2136

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1h
-Q=/cgmt_1/USPTO.spool/US10010720/runat_23092004_154320_25255/app_query.fasta_1.2311
-DB=EST -QPM=fastap -SUFFIX=fst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -DIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10010720 -CCGN_1.1.5750 -runat_23092004_154320_25255 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	2532	23.4	1537	11	BC038700 Mus muscu
2	1909.5	17.7	4705	11	AT321328 Rattus no
3	1814	16.8	3490	11	BC037114 Mus muscu
4	1658.5	15.3	2606	11	AK032505 Mus muscu
5	1594	14.7	1089	11	BM455039 AGENCOURT
6	1565.5	14.5	3623	11	AK052468 Mus muscu
7	1503	13.9	912	12	BT829792 Mus muscu
8	1467.5	13.6	991	13	BK401386 BX401386
9	1428	13.2	1022	12	BM904896 AGENCOURT
10	1380.5	12.8	936	13	BQ988524 AGENCOURT
11	1356	12.5	941	13	BQ956866 AGENCOURT
12	1355.5	12.5	873	13	BQ213010 AGENCOURT
13	1336	12.4	825	9	AU130018 AU130018
14	1328	12.3	829	9	AU131632 AU131632
15	1321	12.2	977	13	BQ058189 AGENCOURT
16	1292.5	12.0	842	14	CD244300 AGENCOURT
17	1277	11.8	805	14	CD644052 AGENCOURT
18	1274	11.8	967	13	BU145705 AGENCOURT
19	1270.5	11.8	918	13	BK415698 BX415698
20	1269	11.7	845	12	BQ976502 AGENCOURT
21	1260	11.7	907	29	CG499521 OST14620
22	1254.5	11.6	830	9	AU132809 AU132809
23	1229	11.4	830	14	CD655395 AGENCOURT
24	1219	11.3	760	13	BX871688 BX871688
25	1209	11.2	721	12	BG378924 UI-R-CV1-
26	1207	11.2	846	14	CFE94281 AGENCOURT
27	1203	11.1	894	13	BQ717639 AGENCOURT
28	1199	11.1	751	14	CB567879 AGENCOURT
29	1197	11.1	905	14	CF553228 AGENCOURT
30	1196	11.1	836	14	CD655715 AGENCOURT
31	1180	10.9	966	12	BT103160 AGENCOURT
32	1175	10.8	894	13	BU178645 AGENCOURT
33	1167.5	10.6	1126	13	BU745817 AGENCOURT
34	1152	10.7	915	12	BG166916 AGENCOURT
35	1148.5	10.6	995	12	BG173180 AGENCOURT
36	1142	10.6	701	13	BO781019 AGENCOURT
37	1137	10.5	698	12	BG722317 AGENCOURT
38	1120	10.4	697	14	CF727167 UI-M-HBO-
39	1117.5	10.3	773	14	CD802866 UI-M-GVO-
40	1115	10.3	676	12	BG768128 AGENCOURT
41	1113	10.3	731	14	CF737595 UI-M-HDO-
42	1108.5	10.2	771	9	AA915514 VZ11h08.x
43	1103.5	10.2	858	9	AL667586 AL667586
44	1103.5	10.2	894	14	CA513151 UI-R-FJO-
45	1091	10.1	750	14	CA511386 UI-R-FJO-

ALIGNMENTS

RESULT 1
BC038700/c 1537 bp mRNA
LOCUS BC038700 Mus musculus, clone IMAGE:1382568, mRNA.
DEFINITION BC038700
ACCESSION BC038700
VERSION BC038700.1 GI:24182477
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1537)
Strausberg,R.


```

QY 265 helysgluAlaGluMetLeuysglYLeuGlnHisProAsnIleValArgPheTYR 285
Db 1167 TCTGTGAAGAGTGTGATGCTCAAGGGGCTGCAGACCCCAACATGCTCCGCTTCTATG 1226
QY 285 spSerTrgIuSerThrVallysglySlyscysIleValIleuValThrGluMet 305
Db 1227 ACTCTGGAAGTGGGTGTGAGGGGTGAGGTTTGATGTGTGTCTCAGAACTCATGA 1286
QY 305 hrSerGlyThrLeuysThrTYrLeuysArgPheIleValMetcysIleysValIleu 325
Db 1287 CCTGGGGACAGCTCAAGACGTAACCTGAGAGATTCGGAATATGAAGCCCGAGTCTTC 1346
QY 325 rgSerTrpCysArgGlnIleLeuysglYLeuGlnPheLeuHisThrArgThrProI 345
Db 1347 AGCGCTGAGTGTGCAGATCTTACGGGAGCTTCACTTCTTACATTCGCGATGCCCCA 1406
QY 345 leIleHisArgAspLeuIlyscysAspAsnIlePheIleThrGlyProThrGlySerVal 365
Db 1407 TCTGTGACCGAGATCTCAAAATGTACAAATGCTTCATTATCCGAGCTTCAGCTGTCA 1466
QY 365 ysIleGlyAspLeuGlyLeuAlaThrLeuysArgAlaSerPheAlaIleValIleG 385
Db 1467 AAATCGAGATCTCGACTGGCCACGCTCAAGCCGCTCTTGGCCAGAGGCTCATGTG 1526
QY 385 lyThrProGluPheMetAlaProGluMetTYrGlnIuIysTYrAspGluSerValAsp 405
Db 1527 GGACCCCGGAATTCATGAGCCCTGAGATGTATGAGAAAGTACGACGAGGCTGTGATG 1586
QY 405 alTYrAlaPheGlyMetCysMetIleuGlnMetAlaThrSerGlyTYrProTYrSerGlu 425
Db 1587 TGTACGCTTTGGCATGTGATGTGAGATGTGTACGTCTGATATCCCTACTCTGAGT 1646
QY 425 ysgIAsnAlaIleGlnIleTYrArgArgValThrSerGlyValIlyscysProAlaSerPhe 445
Db 1647 GTCAGAACGACGACAAATCTACGGAAGGCTCACTTGGGACAAAGCCCAACAGCTTCT 1706
QY 445 spIysValAlaIleProGluValIlysgIuIleIleGlnIyCysIleArgGlnAsnIys 465
Db 1707 ACAAAAGGAAGATGCCAGAGTGAAGAGATCATTTGMAAGCTGCATCCGACGATAGA 1766
QY 465 spGluArgTYrSerIleIysAspLeuLeuAsnHisAlaPhePheGlnIuGlnIy 485
Db 1767 ATAGAGGTTTCACTATCCAGATCTTCTGACCAAGGTTCTTCCGGAGAAAGTGTGTG 1826
QY 485 alArgValGluLeuAlaGluIuAspAspGlyGluIysIleAlaIleIysLeuTYrIleu 505
Db 1827 TGCATGTGAGCTGCGACAGAGAGATGATGAGAGAACCGGCTTCAAACTTGGCTGC 1886
QY 505 rglIleGluAspIleIySlySleuIyGlyTYrIysAspAsnGlnAlaIleGluPheS 525
Db 1887 GCATGAGAGACGACACGACGC---GGGGGGGGCCCAAGGACACAGCCATCGACTTCC 1943
QY 525 ePheAspLeuGluIuAspAspValProGluAspValAlaGlnIuMetcValGluSerGly 545
Db 1944 TGTTCAGCTCGGTGGGATGACGCGAAGAGTGTGCTCAGAGATGAGTCCCTTGGGCT 2003
QY 545 yTYrValCysGlnIyAspHisIysThrMetAlaIlyscysAspArgValSerLeuI 565
Db 2004 TAGTCTGCAAGAGACTACCAAGCCAGTGGCCGTGCAGTTCAGAAACGGGTTCGCGCA 2063
QY 565 leIysArgIysArgGlnIuArgGlnIleuValArgGluGlnIuIySlySglnG 585
Db 2064 TCCAGCGGAAGCGGAGAAAGCTAAGAAAAGCCAGGGAA----- 2101
QY 585 IuGluSerSerLeuysGlnIuValGluGlnSerSerAlaSerGlnThrGlyIleysG 605
Db 2102 -----TTGAGG 2108
QY 605 IuLeuProSerAlaSerThrGlyIlePro---ThrAlaSerThrThrSerAlaSerValS 624
Db 2109 TTTCTCCACCAAGACTCAAGACCTCTCCGCAAGCACTGTGTATGACTCTGCTGCCCCCA 2168
QY 624 eTYrThr---GlnValGluProGluIuProGluAlaAspGlnHisGlnIleuGlnTYrG 643

```

```

Db 2169 GTCCCTTCCCCCAAGAGCCCAAGAGAGCTGAGACCAACCACTCTCTCTTCC 2228
QY 643 IuGlnProSerIle---SerValIleuSerAspGlyThrValAspSerGlyGlnIysers 662
Db 2229 GCCATGCCAGTACTCATTCACATTCATGTATGCGAAGTGAHGTGCTACCTCAGCTTCT 2288
QY 662 eValIlePheThrGluSerArgValSerSerGlnIuThrValSerTYrGlySerGlnHisG 682
Db 2289 CCGGCTTCCGAGATCCCTCAGACCCCTGCTTCAGCCC----- 2326
QY 682 IuGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAlaGlnS 702
Db 2327 -----CCGGGGGAGTGCATTC-----AGCCCCG 2351
QY 702 eGlnProHisGlyValTYrProProSerSerValAlaGlnIyGlnSerGlnIyGlnP 722
Db 2352 CTGAGCCCCAT---CTGTGCTTGCCTCGGTTTTCCTTGTCCATTCACAGGCTGAGCC 2408
QY 722 roSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGlnIyG 742
Db 2409 CTGGCAGT----- 2416
QY 742 IuGlnIyIleGlnIuThrAlaProProGlnIuThrValGlnTYrSerIleuSerGlnT 762
Db 2417 -----GACTTTTCTCTGGGACATGATGTCTGTGTTCACTGATGAGTGC 2462
QY 762 hrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaProGln 782
Db 2463 AGGCTACG-----CCTCAGA 2477
QY 782 alLeuProGlnIuValSerAlaGlyIySglnSerThrGlnIyValSerGlnValAlaPro 802
Db 2478 TGCAGCTCAG----- 2488
QY 802 IuGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerServ 822
Db 2488 ----- 2488
QY 822 alAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnValProS 842
Db 2489 -----CCTCAGTGCATGACGGA----- 2506
QY 842 eSerSerGlyArgHisGluIyArgTYrThrIysArgHisTYrArgIysSerValArgS 862
Db 2507 -----GAAGGGGGGCAATGAGAAATCCAGTGAAGACTCTCGAC 2549
QY 862 eArgSerArgHisGluIysThrSerArgProIysLeuArgIleLeuAsnValSerAsnL 882
Db 2550 GGAGA-----CCTCGATCCCGGCTTGGGTCACAGAGTGTCTCAGACC 2591
QY 882 ysgIAspArgValAlaGluCysGlnIleuGlnThrHisAsnArgIysMetValTh-PheI 902
Db 2592 AGAGGCACAGAGTGTGAGATGCGACCTCAAACTCAACAAGCAAGATGATGACTTTC 2651
QY 902 yspPheAsnLeuAspGlyAspAsnProGluGlnIleAlaThrIleMetcValAsnAsnAsp 922
Db 2652 GATTTGATCTGATGTGGAGACAGCCCAAGAAATGCACCTGCACATGATTATTAATGAAT 2711
QY 922 heIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGluIys 942
Db 2712 TCATTCTGCGCTCGAGAGCAGACGATTCCTAGCCGAGATCCGGGAGATTATTCAGCGAG 2771
QY 942 IaAspGlnMetLeu-----SerGluAspValSerValG 953
Db 2772 TGGAGACCTGCTGAAGAGAGATGCTGTGCTCGAGAGCCACTGAAGATGCGTGAAGC 2831
QY 953 IuProGluGlyAspGlnIyLeuGlnIySerLeuGlnIyIyAspAspTYrGlyPheSerG 973
Db 2832 CTCAGGAAGAGCCACAGACCAAGCTGCTCCAGGCCCATTCAGAT-----GAGG 2882
QY 973 IySerGlnIySleuGlnIyGluPhe-----IySglnProIle----- 985

```


QY 29 SerSerSerAspSerSerValGlyGluIysLeuGlyAlaAlaAlaAlaValThr 48
 Db 121 AGCACCTCG-----AAATACGAGAGCTGGGGCTCC----- 150
 QY 49 GlyArgThrGluIuTrArgArgArgArgIsthrMetAspLysAspSerArgGlyAla 68
 Db 151 -----CATGCCAAACCGAGGCTGACCTGGCTCT 180
 QY 69 AlaAlaThrThrThrThrThrGluHisArgPheArgAspSerValIleCysAspSer 88
 Db 181 GCGGCGC-----ATCTCCGGCTCT 198
 QY 89 AsnAlaThrAlaLeuGluLeuProGlyLeuProLeuSerLeuProGlnProSerIlePro 108
 Db 199 CACCTTACGGGGCGGACCGCTGGGGCGCACCTCTCG-----CCAGAGCGCGCG 249
 QY 109 -----AlaAlaValProGlnSerAlaProProGluProHis 120
 Db 250 CTCTCTGGGAGAGGCTGAGCTCGCGCGCTCTTCGAGACCTGACCGCGCAGCTCAGT 309
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 310 CGA-----TCTGGACTCTGAGCTCTGCTGCTTCAACACCTTCACTCTTCC 357
 QY 141 AlaAlaProGlyGluGlnAlaVal-AlaGlyProAlaProSerThrValProSerSerTh 160
 Db 358 GGAACCCCGGATCCTCCAGACTCCGCTGGCCCCACAGAGAGC-----CCACTTCAAG 411
 QY 160 rSerLysAspArgProValSerGln----- 168
 Db 412 CTCTAAAGAACCCCGGAGGACATGATGGGGGACACTCCCGTGAAGGCTGTGACTC 471
 QY 169 -----ProSerLeuValGlySerLysGluGluProProProAlaArgSerGlySerG 186
 Db 472 TGCATGCTCTGAGCTTACGGGATCT-----TCAGG 501
 QY 186 YGlyGlySerAlaLysGluPro-----GlnGluGluArgSerG 199
 Db 502 AGGTTCAGAGGTCTCCGGGAGCCACCGAGGGTTCCTGATGTCGACCGCGGAGACGCGCGC 561
 QY 199 ngGlnGlnAspArgIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArg 219
 Db 562 GGGAGCGAGAAAGAAAGACACAGAACCCAGGCTGTGGCAACGCTCCGACGCGCGC 621
 QY 219 gPheLeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrIleGlyLe 239
 Db 622 ATACCTCAAGTTGACATTTAGATTGAGAGAGGTTCTTCAAGACCGTGTATGAGAGGCT 681
 QY 239 UAspThrGluThrThrValGluValAlaTyrCysGluLeuGlnAspArgLysLeuThrTy 259
 Db 682 AGACACCGACACACACGAGTAGAGGTGGCTGTGTGAGTGCACACCGGAAATCTCTCTG 741
 QY 259 sSerGlnArgGlnArgPheLysGlnGluAlaGluMetLeuLysGlyLeuGlnHisProAs 279
 Db 742 GCGCGAGCGCGAGCGGTTCTTCGAGAGAGTTCAGATTCAGAGGCGCTCAGACCCCA 801
 QY 279 nIleValAlaArgPheTyAspSerTrpGluSerThrValLysGly-LysLysCysIleVal 299
 Db 802 CATCGCGCTCTTACGACTCTCGAGAGTCCGCTGAGGGCGCAGG----- 848
 QY 299 euValThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysVal 319
 Db 849 -----TACCTGAGCGCGTTCCGAGAA 870
 QY 319 eLlyIleLysValLeuArgSerTrpCysArgGlnIleLeuLysGlyLeuGlnPheLeuH 339
 Db 871 TGAAGCGCGGAGTCTTCAGCGCTGAGTGCACAGATTCTACGAGGACTTCACTTCTTC 930
 QY 339 iSthrArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThng 359
 Db 931 ATTCCGAGGTGCCCGCATCTCTGACACGAGATCTCAAGTGTGACAAACGCTTCACTTACG 990

QY 359 lYProThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSer 379
 Db 991 GCCCTTCAGGCTCTGTCAAAATCGAGACTCGGACTTGGCCACGCTCAAGCGCGCTCTCT 1050
 QY 379 heaLalysSerValIleGlyThrProGluPheMetAlaProGluMetTyrTrpGluLysTr 399
 Db 1051 TTGCCAAGAGGTATTTGGAGCCCGGAGATTCAATGAGCTCTGAGATGTACAGAGAAAGT 1110
 QY 399 YrAspGluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerg 419
 Db 1111 ACAGACGAGCTGTGATGTGATACGCTTTGGCATGTGATCTGGAGATGGCTACGTCTG 1170
 QY 419 luyTrProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyV 439
 Db 1171 AGATTCCTTACTCTCGAGTGTGAGATTCAGATTCAGACAAATCTTCCGAGAGTCACTTCGGGCA 1230
 QY 439 aLysProAlaSerPheAspLysValAlaIleProGluValLysGluIleGluGlyC 459
 Db 1231 CAAGGCCAAGAGCTTCTTACAGGTGAGAGATGCCAGAGTGAAGAGATCATTTGAAGGCT 1290
 QY 459 YsIleArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPheP 479
 Db 1291 GCATCCGACAGGATTAAGAAAGAGAGGTTCAACATCCAGATCTCTGGCCCATGCTTCT 1350
 QY 479 heGlnGluGluThrGlyValArgValGluLeuAlaGluGluAspArgGlyGluLysIleA 499
 Db 1351 TCCGCGAGGAGACAGAGTGTGCACGTGAGCTGGCAGAGAAATGATGAGAGAGACCGG 1410
 QY 499 lalIleLysLeuTrpLeuArgIleGluAspIleLysLysLysGlyLysTyrLysAspA 519
 Db 1411 GCCTCAAGCTCTGCTGCTGCGCATGAGAGATCGCGGCGC---GGGGGGCGTCCACAGACA 1467
 QY 519 snGluAlaIleGluPheSerPheAspLeuGluLysArgAspValProGluAspValAlaGln 539
 Db 1468 ACCAGGCATCGAGTTCCTGTTCACTCGCGCGGAGCGCTGGGAGAGAGTGGCTCAGG 1527
 QY 539 lmetValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleL 559
 Db 1528 AGATGGTAGCTTGGGCTTAGTGTGAGCTGACTACCAAGCCAGAGGCGCCGCGAGTTC 1587
 QY 559 YsAspArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGlu 579
 Db 1588 GAGAACGGGTTGTGCTATTCAGCGGAGAGGAGAAAGCTTGAAAGACTGAGGAA--- 1643
 QY 579 lngLulysLysGlnGluGluSerSerLeuLysGlnGlnValGluGlnSerSerAlas 599
 Db 1643 ----- 1643
 QY 599 ergLlnThrGlyLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThr 619
 Db 1644 -----TTGGAGTTCCTCCACCAAGACTCA---GAGCTCTCCAGCAACGTGTG 1689
 QY 619 hSerAlaSerValSerThrGlnVal-----GluProGluGluProGluAlaAspG 636
 Db 1690 CTCTGCTCGCGGTCCTCCCGCAGTGCCTTCCCGCAGAGCTGAGAGCCAGAGGCTGAGC 1749
 QY 636 lnhIleGlnLeuGlnTyrGlnGlnProSerIle---SerValLeuSerAspGlyThrV 655
 Db 1750 AGCATCAATCTCTCTCTCCCGCATGCCAGTCACTCATCACTCATCATCTGATTGGAGA 1809
 QY 655 aLAspSerGlyGlnGlySerSerValPheThrGluSerArgValSerSerGlnGlnThv 675
 Db 1810 CTGATGGCTACTCTCAAGTTCCTCCGCGCTCTCGAGATCCAGACCTGCCCTTCAGCCC- 1868
 QY 675 aLserTyrGlySerGlnHisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleP 695
 Db 1869 -----CCTGGGGGGTTTAC 1881
 QY 695 rOserThrValGlnAlaGlnSerGlnProHisGlyValTyrProProSerSerValAlaG 715
 Db 1882 CATTCAGC----- 1889
 QY 715 lngLysGlnSerGlnGlnGlnProSerSerSerSerLeuThrGlyValSerSerSerGlnP 735

QY	1039	SeThrhzLaGInSerProGlyMeTAnLeuSerHisSerLaSerSerLeuSerLeuGln	1058
Db	744	TCACATCTCAAGGCCCTCGAAGAAGAACTTATCTCACTCTGCTTCACTCTCCAGCTCCGAG	803
QY	1059	GlnAlaPheSerGluLeuArgAlaGlyMeTThrGluGlyProAsnThrAlaProPro	1078
Db	804	CAGGCTTTTCTGTAACGTGAACAAATGTCATATGATGTAAGAACCCATATACGACCTCCA	863
QY	1079	AsnPheSerHis--ThnGlyProThrPheProValValProProPheLeuSerSerIle	1097
Db	864	AATTTTATATACATGCGGAGCAACAACTTT-----TCTCCCTTTTGGCTACACTT	914
QY	1098	AlaGlyValProThrThrAlaAlaAlaThr-----AlaProValProAlaThrSerSer	1115
Db	915	GCTGGTGCACAACTGATAGACATCCACCCCTGATCATGCTCCCTATTAACAACAGT	974
QY	1116	ProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGluGln	1135
Db	975	CCTCTTATATGACATTTCCACATCACTGATGATGACAGTCAAGCTGCACCTGCCCACTGA	1034
QY	1136	GlyIleAlaGlyVal---AlaThrSerThrGlyValValThrSerGlyLeuProIle	1154
Db	1035	GGCATCGGGAGATTACACATACAGACAGAGTGTGTACTTCAAGGTGCTTCACACACA	1094
QY	1155	ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla	1174
Db	1095	ATGTCGTGTGTGAGTCAACCAACTCTTCCAGTGGGTTCMAAGTTCTTACAGTACTGCA	1154
QY	1175	ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle	1194
Db	1155	GTTTGATCATGTGTCCACA-----CCTCCAGCCAGTTCAGGGTTCACATCTGGGAGC	1208
QY	1195	ValValSerSerThrAlaLeuTyProSerValThrValSerAlaThrSerAlaSerAla	1214
Db	1209	ATTGCTTCAGATGACAGGAGATTTCCTCTCGAACAATTTCTTCAACATCAGACTACTACA	1268
QY	1215	GlyGlySerThrThrAlaThrProGlyProLysProProAlaValAlaSerGlnGlnAlaAla	1234
Db	1269	ATGGGTAGTGGTGGGCGCCCAAGATGCTAAGCCTCCCACTGATTATTTGCAACAGATGACA	1328
QY	1235	GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer	1254
Db	1329	AGCAATATCGCTGGGGTTCCTCCATGATGATCATCAAGTCTTACACACATCCATTCGAGGC	1388
QY	1255	ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrLeuAlaGlu	1274
Db	1389	ATGGGTTCACACACCAATCCCTTCCACTTGTATGACAGACCTCTGCTCGTACTGTAGCTGAA	1448
QY	1275	ThrValValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu	1294
Db	1449	ACAAATGGTGTATGAGCACACATCACTTAACAAGGATGCTACAGATGACAGACAGAGATG	1508
QY	1295	AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyx	1314
Db	1509	GGTTTGTCTTTTGTGGCCCATCTCTTCTCTCTCTCGAACAAGCTGTGTCTACTTGT	1568
QY	1315	IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro	1334
Db	1569	GTTTTCACAGCGATATGGTTCATCTTTGTGATCATGATACAGGATGTTGTATACAGCT	1628
QY	1335	IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle	1354
Db	1629	GGCTTCGCCCAACACAGTATACCTTCTTACACCTTTATACCCCAAGTACTTAATTC	1688
QY	1355	ProProLeuValGlnProValAlaSerValProAlaValGlnGlnThrLeuIleHisSer	1374
Db	1689	CCACCTTGTATACACTTGTATTATGTGCTGCTGTCATACGAAACACTCATTCACAGT	1748
QY	1375	GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer	1394
Db	1749	CAGCTTCACACAGCTCTGCTTCCCAACAGCCTTCACACACACTGTCTCGAATATGATGCT	1808
QY	1395	AspThrGlnProLysAlaProGlyIleAspAspIleLysThrLeuGlnGlnLysLeuArg	1414

Percent Similarity:	96.69%	Conservative:	0
Best Local Similarity:	96.69%	Mismatches:	8
Query Match:	14.74%	Indels:	4
DB:	12	Gaps:	0
US-10-010-720-14 (1-2136) x BM455039 (1-1089)			
QY	1808	PheAsnSerSerTyMetSerSerAspAsnGluSerAspGluAspLeuLys	1827
DB	2	TTTAACTCTCTTACATAGTAGGACCAATGAGTACAGTATCGAAGATGAAGACTTAAG	61
QY	1828	LeuGluLeuArgArgLeuArgAspLysHisLeuLysGluLeuGlnAspLeuGlnSerArg	1847
DB	62	TTAAGGCTGCGACAGCTACGAGATTAACATCTCAAGAGATTGAGCACTGACAGTGC	121
QY	1848	GlnLysHisGluLeuLeuSerLeuTyThrLysLeuGlyLysValProProAlaValIle	1867
DB	122	CAGAAAGCATGAATTTGAATCTTTGTATTCGAACTGGGCAAGGAGGCCCTGCTGTATT	181
QY	1868	LeuProProAlaAlaProLeuSerGlyArgArgArgProThrLysSerLysGlySer	1887
DB	182	ATTCCCCAGCTGCTCCCTTTCAGGGAGAAAGACGAGCCACTAAAGCAAGGACGC	241
QY	1888	LysSerSerArgSerSerSerSerLeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSer	1907
DB	242	AAATCTAGTGAAGACAGTCTCTTGGGGAATMAAGCCCCCACTTTCAGGTAACTGTCT	301
QY	1908	GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLeuHisProProGlyAsnIle	1927
DB	302	GGTCAGAGTGCAGCTTACGCTTTCACCCCGACGACCCCTCCCTCCGGAACATC	361
QY	1928	ProGlnSerGlyGlnAsnGlnLeuLeuGlnProLeuLysProSerProSerSerAspAsn	1947
DB	362	CCAGAGTCCCGGAGATCAGCTGTCTACGCCCCCTTAAGCATCTCCCTCCAGTGAAC	421
QY	1948	LeuTySerAlaPheThrSerAspGlyAlaIleSerValProSerLeuSerAlaProGly	1967
DB	422	CTTATATTCACCTTCACCACTGATGAGGCCATTCATTCACCAAGCCCTTCTGCTCCAG	481
QY	1968	GlnGlyThrSerSerThrAsnThrValGlyAlaThrValAsnSerGlnAlaGlnAla	1987
DB	482	CAAGGAACCGACGACCAACACTGTTGGGSCAACAGTGAACGCCAAGCCGCGCAAGCT	541
QY	1988	GlnProProAlaMetThrSerSerArgLysGlyThrPheThrAspAspLeuHisLysLeu	2007
DB	542	CAGCCCTCCGCTGATGCTCCACAGAGAGGACATTCACATAGACTTGCACAAAGTTG	601
QY	2008	ValAspAsnThrAlaArgAspAlaMetAsnLeuSerGlyArgArgLysGlyHis	2027
DB	602	GTAAGCAATTTGGGCGGAGATGCAATCTCTCAGGAGGAGAGAGCAAGCAAGGGCAC	661
QY	2028	MetAsnTyGlnGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCysIle	2047
DB	662	ATGAATTAAGAGGCGCTCGAAAGGCAAGAGAGTTCTTGCACCTTGAGCAACTGTCTATC	721
QY	2048	SerMetThrSerAsnLeuGlyLysSerAlaProIleSerAlaAlaSerAlaThrSerLeu	2067
DB	722	TCCATGACCTCGAACTGAGTGGCTGCTGCCCATCTTGCAGCATCACTACCTCTCTA	781
QY	2068	GlyHisPheThrLysSerMetCysProProGlnGlnTyGlyPheProAlaThrProPhe	2087
DB	782	GGTGACCTTACCAAGCTATGAGCCGCCCAAGCAGTATGGCTTCCACTAC-CCATT	840
QY	2088	GlyAlaGlnTrp-SerGlyThrGlyGlyProAlaProGln-ProLeuGlyGlnPheGlnP	2107
DB	841	GCGCGCTCATAGGAGTGGGAGCGAGTGGCCCAACCAACGAGCCACTTGGCCAGTTTCAA	900
QY	2107	rovalGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerIleSera	2127
DB	901	CTGTGGGAGTGGCCCTCTTTCAGAAATTTCAATCAGCAATTTGCAAGAAATCCATCAGCA	960
QY	2127	sn-ProProGlySerAsnLeuArgThrThr	2136

DB	961	ACCCCCCAGGCTTCAACTGCGGACACC	990
RESULT 6	AK052468		
LOCUS			
DEFINITION	AK052468	3623 bp	mRNA
ACCESSION	AK052468		linear
VERSION	AK052468.1	GI:26095199	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1	Carninci, P. and Hayashizaki, Y.	
AUTHORS	1	High-efficiency full-length cDNA cloning	
TITLE	1	Meth. Enzymol. 303, 19-44 (1999)	
JOURNAL	1	99279253	
MEDLINE	1	10349636	
PUBMED	1		
AUTHORS	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibata, K.,	
TITLE	2	Normalizing and subcloning of cap-trapper-selected cDNAs to	
JOURNAL	2	prepare full-length cDNA libraries for rapid discovery of new genes	
MEDLINE	2	Genome Res. 10 (10), 1617-1630 (2000)	
PUBMED	2	20499374	
AUTHORS	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,	
TITLE	3	Komno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,	
JOURNAL	3	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,	
MEDLINE	3	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,	
PUBMED	3	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,	
AUTHORS	3	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, O.,	
TITLE	3	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
JOURNAL	3	RIKEN integrated sequence analysis (RISA) system -384-format	
MEDLINE	3	Genome Res. 10 (11), 1757-1771 (2000)	
PUBMED	3	20530913	
AUTHORS	4	11076861	
TITLE	4	The RIKEN Genome Exploration Research Group Phase II Team and the	
JOURNAL	4	FANTOM Consortium.	
MEDLINE	4	Functional annotation of a full-length mouse cDNA collection	
PUBMED	4	Nature 409, 685-690 (2001)	
AUTHORS	5		
TITLE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research	
JOURNAL	5	Group Phase I & II Team.	
MEDLINE	5	Analysis of the mouse transcriptome based on functional annotation	
PUBMED	5	of 60,770 full-length cDNAs	
AUTHORS	6	Nature 420, 563-573 (2002)	
TITLE	6	(bases 1 to 3623)	
JOURNAL	6	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,	
MEDLINE	6	Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,	
PUBMED	6	Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, I., Hirozane, T.,	
AUTHORS	6	Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,	
TITLE	6	Kato, H., Kawai, D., Kojima, Y., Kondo, S., Komno, H., Konda, M.,	
JOURNAL	6	Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,	
MEDLINE	6	Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N.,	
PUBMED	6	Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,	
AUTHORS	6	Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,	
TITLE	6	Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akihara, S.,	
JOURNAL	6	Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,	
MEDLINE	6	Muramatsu, M. and Hayashizaki, Y.	
PUBMED	6	Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of	
AUTHORS	6	Physical and Chemical Research (RIKEN), Laboratory for Genome	
TITLE	6	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),	
JOURNAL	6	RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsunumi-ku, Yokohama,	
MEDLINE	6	Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,	

COMMENT

URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>,
 URL: <http://fantom.gsc.riken.go.jp/>,
 Location/Qualifiers

FEATURES

source

1. 3623

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM_DB:D430030M17"

/db_xref="MGI:2421908"

/db_xref="taxon:10090"

/clone="D430030M17"

/tissue_type="lung"

/clone_id="RIKEN full-length enriched mouse cDNA library"

misc_feature

1. 3623

/note="hypothetical Eukaryotic protein kinase/Glutamine-rich region containing protein (InterPro|IPR00719, PROSITE|PS50322, evidence: InterPro)"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

US-10-010-720-14 (1-2136) x AK052468 (1-3623)

	Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:
427	AspAlaAlaGlnIleTyrArgArgValThrSer	5.12e-69	Length: 3623	Matches: 1565.50	Conservative: 82.15%
3	AATGCTGGCAGATCTACCGTGAAGTGAAGGTTTATGAAAGTGCAGCCACTGAGAA	82.15%	Mismatches: 12	Conservative: 82.15%	Mismatches: 20
437	-----	14.48%	Indels: 53	Indels: 53	Gaps: 5
63	AAGAGTCATCACCAGGCTTCCCTCAGCTGAAAGTTAGCTAAGATAAAGCTGCGCTGCC	12	-----	437	
438	-----	437			
123	AAGATATGCTACTGCTGCAACAGTGGCAGATGTTATGAGGAGGAGTGAAGCCAGCAGT	182			
444	PheAspLysValAlaIleProGluValIleGluIleIleGluGlyCysIleArgGlnAsn	463			
183	TTTGACAAAGTAGAATCTCTGAAAGTAAAGAAATTTAGAAAGATGTTATCCGACAAAC	242			
464	LysAspGluArgTyrSerIleLysAspLeuLeuAsnIleApeheGlnGluThr	483			
243	AAAGATGAAGATATTTCCATCAAGACCTTTTGAACATGCTTTTCCAGAGGAAACA	302			
484	GlyValArgValGluLeuAlaGluGluAspGlyGluIleAlaIleLysLeuTyr	503			
303	GGGGTACGGGTGAATTAGCAGAAAGATGATGAGAAAGATAGCTATTAATTTATGG	362			
504	LeuArgIleGluAspIleLysLysLeuLysGlyIleTyrLysAspAsnGluAlaIleGlu	523			
363	CTGGGTATTTGAAGATATTGAAGAAATTAAGGAAATTAAGCAATGAAGCTATTGAG	422			
524	PheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMetValGluSer	543			
423	TTTTCTTTGACTTGGAGAGGAGTACCAAGATGTTGCTCAAGAAATGTTGAGCT	482			
544	GlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAspArgValSer	563			
483	GATTATGCTGTGAAGGATGATCACAAGCAGTGGTAAAGCTATCAAGAGATAGGCTGTC	542			

QY	564	LeuIleLysArgLysArgGluGlnArgGlnLeuValAlaArgGluGluGlnGluLysLys	583
DB	543	TTAATTAGAGAAACAGAGACAGCGCAGTGGTTCAGAGAGACAAAAAGAAAA	602
QY	584	GlnGluGluSerSerLeuLysGlnGlnValGlu---GlnSerSerAlaSerGlnThrGly	602
DB	603	CAGAGAGAGAGAGGTTTCAACAGCAATGAAACAAAGCAAGAGGCTTTCCAGGACGAA	662
QY	603	IleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrsAlaSer	622
DB	663	ATCCAGCAGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	722
QY	623	ValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeuGlnTyr	642
DB	723	GTTTCTACACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	782
QY	643	GlnGlnProSerIleSerValLeuSerAspGlyTyrValAlaAspSerGlyGlnGlySer	662
DB	783	CAGCAACCTAGTATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	842
QY	663	ValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGlnHisGln	682
DB	843	GTAATCAGAAATCCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	902
QY	683	GlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAlaGlnSer	702
DB	903	CAGCAGATTTACTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	962
QY	703	GlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGlnPro	722
DB	963	CAGCCCAATGAGGATGATACCTTCAAGTATGAGCAGAGGAGCAGAGGAGGAGGAGG	1022
QY	723	SerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGlnGln	742
DB	1023	---TCAAGTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1076
QY	743	GlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSerGlnThr	762
DB	1077	CAGGAAATCAGCAATCTGCTCTCTCAAGAGCAGTACAGTATCACTTCCACAGCA	1136
QY	763	SerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaProGlnVal	782
DB	1137	GCATCTTCCAGTGAAGTACTCTGCAAGCCAGTGAAGTGAAGTGAAGTGAAGTGAAG	1178
QY	783	LeuProGlnValSerAlaGlyLysGln	791
DB	1179	---CTCAAGTGTCACTGGAACACAG	1202
RESULT 7			
BI829792			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
NIH-MGC http://mgs.nci.nih.gov/ . CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at:			

DB: 13 Gaps: 0

US-10-010-720-14 (1-2136) x BX401386 (1-991)

QY 1182 SerProSerLeuGlnValProThSerThSerGluValValSerSerThraLeu 1201

DB 52 TCCCGGAGATCTTCAAGTCCCAATCCATCATCTGAGATGTTGTTCTTAGTACACAGC 111

QY 1202 TyrProSerValThValSerLarThSerLarSerAlaGlyGlySerThraLarPro 1221

DB 112 TATCTCTGAGTACAGTTTCAAGCACTTACGCTTGCAGGGGAGAGTACTCTACCCCA 171

QY 1222 GlyProLysProProAlaValValSerGlnGlnAlaAlaGlySerThraValGlyAla 1241

DB 172 GGTCTTACGCTTCCAGCTTATGATCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 231

QY 1242 ThrLeuThSerValSerThThThThThThThThThThThThThThThThThThTh 1261

DB 232 ACATTACATCAAGTTCTTACACACACTTCTTCCCAAGCAGCAGCAGCAGCAGCAGC 291

QY 1262 GlnLeuSerSerSerThSerThThThThThThThThThThThThThThThThThTh 1281

DB 292 CAGCTTACAGCAGCAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 351

QY 1282 SerLeuAspLysThSerThSerThSerThThThThThThThThThThThThThTh 1301

DB 352 TCCTTACATGATGACATCTTCAATGACATGACATGACATGACATGACATGACATGAC 411

QY 1302 SerSerSerSerSerProGlyAlaGlyValSerSerTyrlleSerGlnProGlyGly 1321

DB 412 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 471

QY 1322 HisProLeuValLleProSerValLleAlaSerThProLleLeuProGlnAlaAla 1341

DB 472 CATCTTGTGTCATCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 531

QY 1342 ProThSerThProLeuLeuProGlnValProSerLleProProLeuValGlnProVal 1361

DB 532 CCACTTCTTACACCTTATTAATCCCAAGTACCTAGTACCTAGTACCTAGTACCTAGT 591

QY 1362 AlaAsnValProAlaValGlnGlnThLeuLleHisSerGlnProGlnProAlaLeu 1381

DB 592 GCCAATGTGCTGCTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 651

QY 1382 ProAsnGlnProHisThThHsCysProGlnValAlaSerSerAspThGlnProLysAla 1401

DB 652 CCCAACCAAGCCCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 711

QY 1402 GlyLleAspAspLleLysThThLeuGlnGlnLysLeuArgSerLeuPheSerGluHisSer 1421

DB 712 GGAATTAATGACATTAAGAAATTTAGAAAGAAAGCTGGGCTCTCTCTTCAAGTAAACAGC 771

QY 1422 SerSerGlyAlaGlnHisAlaSerValSerLeuGlnThSerLeuValLleGluSerTh 1441

DB 772 TCATCTGAGCTGAGCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 831

QY 1442 ValThProGlyLleProThThThThThThThThThThThThThThThThThThTh 1461

DB 832 GTCAACACAGGACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 891

QY 1462 SerThCysLeuProProThThThThThThThThThThThThThThThThThThTh 1481

DB 892 AGTACTGTTTCCACCAACCAATTAACCAAGTGTGTTTCCAGTTTACACCA 951

QY 1482 ValValThProGlyGlnValSerThProValSerThThThThThThThThThTh 1495

DB 952 GTGTCTACACT-GGGCAAGTTT-TTCCGCTCAGCAGCTACT 991

RESULT 9

BM904896 1022 bp mRNA linear EST 12-MAR-2002

LOCUS BM904896

DEFINITION AGNCOURT 6699439 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557718

5', mRNA sequence.

ACCESSION BM904896

VERSION BM904896.1 GI:19355275

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC http://mgi.nci.nih.gov/

TITLE NIH-MGC Consortium of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cga@bbs-remail.nih.gov Tissue Procurement: ATCC/DCTD/DTF CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Plate: L14M12279 Row: 1 Column: 15 High quality sequence start: 3 High quality sequence stop: 715. Location/Qualifiers

FEATURES

source

1..1022

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5557718"

/issue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_1ib="NIH_MGC_72"

/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 9,54e-63 Length: 1022

Score: 1428.00 Matches: 308

Percent Similarity: 91.55% Conservative: 6

Best Local Similarity: 89.80% Mismatches: 16

Query Match: 13,214 Indels: 13

DB: 12 Gaps: 4

US-10-010-720-14 (1-2136) x BM904896 (1-1022)

QY 1302 SerSerSerSerSerProGlyAlaGlyValSerSerTyrlleSerGlnProGlyGly 1321

DB 9 TCGTCGACCAAGCGCTCGGAGCAGGAGTGTCTAGTTATATTTCTCAGCTGTGGGCTG 68

QY 1322 HisProLeuValLleProSerValLleAlaSerThProLleLeuProGlnAlaAla 1341

DB 69 CATCTTGTGTCATCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 128

QY 1342 ProThSerThProLeuLeuProGlnValProSerLleProProLeuValGlnProVal 1361

DB 129 CCACTTCTTACACCTTATTAATCCCAAGTACCTAGTATCCCACTTGTGTAAGCTGT 188

QY 1362 AlaAsnValProAlaValGlnGlnThLeuLleHisSerGlnProGlnProAlaLeu 1381

DB 189 GCCAATGTGCTGCTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 248

QY 1382 ProAsnGlnProHisThThHsCysProGlnValAlaSerSerAspThGlnProLysAla 1401

DB 249 CCCAACCAAGCCCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 308

QY 1402 GlyLleAspAspLleLysThThLeuGlnGlnLysLeuArgSerLeuPheSerGluHisSer 1421

DB 309 GGAATTAATGACATTAAGAAATTTAGAAAGAAAGCTGGGCTCTCTCTTCAAGTAAACAGC 368

QY 1422 SerSerGlyAlaGlnHisAlaSerValSerLeuGlnThSerLeuValLleGluSerTh 1441


```

Db      369  TCATCTGAGGTCACACATGCTCTCTCTCAGAGACCTCAGTACTAGAGGACCT 428
QY      1442  ValThrProGlyIleProThrThraIaIaIaProSerIyLeuLeuThrSerThrThr 1461
Db      429  GTCAACACAGGATTCACCACTACTCTGTTGACCAAGAACTCTGACTTCTACACA 488
QY      1462  SerThrCysLeuProProThrAsnLeuProLeuGlyThrValaIaIaLeuProValThrPro 1481
Db      489  AGTACTCTCTTACCAACCAACATTTACCACTAGAACAGTTGCTTCCAGTTACACA 548
QY      1482  ValValThrProGlyIleValSerThrProValSerThrThrThrSerGlyValIySerPro 1501
Db      549  GTGGTCAACCTGGGCAAGTTTCACTCCAGTCACTACTACTACTAGAGAGTGAACCT 608
QY      1502  GlyThrIaIaProSerIyProProLeuThrIyIaIaIaProValIleProValGlyThrGlu 1521
Db      609  GGAATGCTGCTCCCTCCAGCCACCTCTAAGGCTCCGGTGTGCCAGTGGTACTGAA 668
QY      1522  LeuProIaIaGlyThrLeuProSerGlyIleLeuProProPheProGlyProSerLeuThr 1541
Db      669  CTTCCACACAGGATCTTACCAACGAGAGAGCTGCCACCTTTCCAGAGACCTTCTAAC 728
QY      1542  GluSerGlnGln-ProLeuGluAspLeuAspIaIaGlnLeuArgThrThrLeuSerProGly 1561
Db      729  CAGTCCACGACACCTCTAGAGAGCTTGATGCTCAATGAGAAACACTTACTCTCAGA 788
QY      1561  U***IleThrValThr-SerIaIaIaGlyProValSerMetIaIa-AlaProThrAla-I 1580
Db      789  GATGATCACAGTACCTTCTGCGGNTGCTCTGTCCTCAATGCGGGCTCCACAGCCAA 848
QY      1580  eThrGluAla-GlyThrGlnProGlnIyGlyVal---SerGlnValIyGlnIyGlyProV 1599
Db      849  CACAAAACGACGAGAACCCCTCAGAAAGTGTCTTCTCAAGAACGAAAGAGCCCTG 908
QY      1599  alLeuAla-ThrSerSer-GlyIaIaGlyValPheIyMetGlyY-----ArgPheG 1615
Db      909  TCTTACCACTAGTATTAAAGAGTGGGGTTTTAAAGAGGAGCAAAATTTCCGGTTCC 968
QY      1615  nValSerValaIaIaAspGlyIaIaGlnIyGlnIyGly---LysAsnIySerGln 1632
Db      969  T---GGTTGACAAACGAGGCCCCCAAAAGGGGTAAAGATTAATGTCAAA 1020

RESULT 10
LOCUS   BQ898524          936 bp      mRNA      linear      EST 16-AUG-2002
DEFINITION
AGENCOURT 8728264 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6339300
5', mRNA sequence.
ACCESSION BQ898524
VERSION   BQ898524.1 GI:22290538
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 936)
NIH-MGC http://mgi.nci.nih.gov/
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bhs-neml.nih.gov
Tissue Procurement: ATCC
cDNA library Preparation: Rubin Laboratory
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LDCM253 row: 0 column: 13
High quality sequence stop: 656.
Location/Qualifiers
1..936
/organism="Homo sapiens"

```

```

/mo1_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6339300"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 47"
/note="Organ: brain; Vector: pOTB7, Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      2,196-60      Length:      936
Score:          1380.50      Matches:      293
Percent Similarity: 94.58%      Conservative: 2
Best Local Similarity: 93.91%      Mismatches: 12
Query Match:    12.77%      Indels:      5
DB:              13          Gaps:        3

US-10-010-720-14 (1-2136) x BQ898524 (1-936)
QY      915  ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln 934
Db      1  ACAATTATGTTGAACAATGACTTATTATCTAGCAATAGAGAGAGTGCTTTGTGATCA 60
QY      935  ValArgGluIleIleGluIyValaIaAspGluMetLeuSerGluAspValSerValGluPro 954
Db      61  GTGCGAAGAAATTAATGAAGAGCTGAGTGAATGCTCAGTAGATGTCAGTGTGGAACA 120
QY      955  GluGlyAspGlnIyLeuGluSerLeuGlnIyLysAspAspTyrGlyPheSerGlySer 974
Db      121  GAGGATGATAGAGGATTTGAGAGATCTACAGAGAAAGATGACTATGACTTTTCAAGTTCT 180
QY      975  GlnIyLeuGlnIyGluPheIyGlnProIleProIaIaSerSerMetProGlnIle 994
Db      181  CAGAAATTTGAGAGAGAGTGAACAACCAATTCTCGGCTTCCATGCCACGCAATA 240
QY      995  GlyIleProThrSerSerLeuThrGlnValValHisSerIaIaGlyArgPheIleVal 1014
Db      241  GGCATTCTTACCAAGTCTTAACTCAAGTGTTCATTTCCGGGAGAGCGGTTTATATG 300
QY      1015  SerProValProGluSerArgLeuArgGluSerIyValPheProSerGluIleThrAsp 1034
Db      301  AGTCCCTGTCACAGAAAGCCGATTACAGAAATCAAAAGTTTCCCGAGTGAATTAACAGAT 360
QY      1035  ThrValaIaIaSerThrIaIaGlnSerProGlyMetAsnLeuSerHisSerIaIaSerSer 1054
Db      361  ACAGTTGCTGCTCTTACAGCTCAGAGCCCTGGAATGAACCTTGTCTCACTGTGATCATCC 420
QY      1055  LeuSerLeuGlnIaIaPheSerGluLeuArgAlaGlnMetThrGlnIyGlyProAsn 1074
Db      421  CTTAGCTTACAAAGGCTTCTGAACTTAACGCGCCAAATGACAAAGAGACCAAC 480
QY      1075  ThrIaIaProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu 1094
Db      481  ACAGACCTCCCAAACTTTAGATACAGACCAACATTTCCAGTAGTACTCTTTCTTA 540
QY      1095  SerSerIleAlaGlyValProThrThrIaIaIaIaIaIaProValProIaIaThrSer 1114
Db      541  AGTACATTTGCTGGAGTCCCAACACAGAGACGACGACGACGCTCCCTCAACAGC 600
QY      1115  SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db      601  AGCCCTCTTAATGACATTTCCACATCAGTAATTCAGTGTGAGTTACAGTCCCACTGAA 660
QY      1135  GluGlyIleAlaGlyValaIaThrSerThrGlyValaIaIaIaIaIaIaIaIaIaIaIa 1154

```

Db 661 GAGGGATGCTGAGTTGCCACAGCAGGTGTGTAACTTCAGTGTCTCCCATTA 720

QY 1155 ProProValSerGluSerProValIleuSerSerValIleSerIleThrIleProAla 1174

Db 721 CCACCGTGTCTGAAATCACCAGTACTTTCACGAGTTCAGATTCACAAATACCTGCA 780

QY 1175 ValIleSerIleSerThrThrSer-ProSerLeuGluValProThrSerThrSer---G1 1193

Db 781 GTTGTCTCAATATCTACTACATCCCGGCTACTTCAGTCCCAATCCACATCTGGAG 840

QY 1193 uIleValIleSerSerThrAlaIleu-TyrPro---SerValThrValSerIleThrSera 1212

Db 841 ATCGGTGTCTTCTAGTACAGCACTGGTATCTTCAATACAGTTTTCAGCAAACTTCAN 900

QY 1212 Ia---SerIleGlyIleSerThrAlaThrPro 1221

Db 901 NCCCTTTGAGAGGGGGGAGTACTGCTTACCC 932

RESULT 11

LOCUS BQ956866 941 bp mRNA linear EST 21-AUG-2002

DEFINITION AGENCOURT 8810997 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6429116

ACCESSION BQ956866

VERSION BQ956866.1 GI:22372344

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS NIH-MGC http://mgs.nci.nih.gov/

TITLE 1 (bases 1 to 941)

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: LCM2616 row: m column: 21

High quality sequence start: 4

High quality sequence stop: 623.

FEATURES

source

1..941

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6429116"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 101"

/note="Organ: Lung; Vector: pOTB7; Site:1: EcoRI; Site:2: XhoI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 3,866-59 Length: 941

Score: 1356.00 Matches: 288

Percent Similarity: 93.65% Conservative: 6

Best Local Similarity: 91.72% Mismatches: 13

Query Match: 12.54% Indels: 8

DB: 13 Gaps: 2

US-10-010-720-14 (1-2136) x BQ956866 (1-941)

QY 1582 GluAlaGlyThrGlnProGlnIleuSerGlyValSerGlnValIleGlyIleProValIleAla 1601

Db 1 GAGGAGGAGACAGACCTCCAGAGGCTTTTCTCAAGTCAAAGAGGCTCTCTAGCA 60

QY 1602 ThrSerSerGlyAlaGlyValPheIleSerGlyValArgPheGlnValSerValAlaIleAsp 1621

Db 61 ACTAGTTCAGAGGCTGTGTGTTTAAAGATGGAGCAATTCAGGTTCTGTTCAGCAGAC 120

QY 1622 GlyAlaGlnIleuSerGlnIleValSerGlnIleuSerGlnIleuSerValIleSerPheGlnSer 1641

Db 121 GGTTGCCGAAAGAGGATTAATATAGTCAGAGATGCAAACTCTCTTCAATTTGATTC 180

QY 1642 SerThrSerGluSerSerValIleuSerSerSerSerProGlnSerThrIleuValIlePro 1661

Db 181 AGCACTCGAGTCTCTCACTGCTATCAATAGTACGAGAGTACCTTGATGAACCA 240

QY 1662 GluProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHisIleVal 1681

Db 241 GAGCCGATGCGATTAACCATCCCTGTATCTTTCAGATGTGCCAGAGGTGCCCAAA 300

QY 1682 ThrThrAlaSerGlnAlaIleuSerSerThrGlyIleProThrIleValGlyArgPheGln 1701

Db 301 ACTACTGCTCTCAGAGCAAGATCAGACTGGGAGCTTCCAGAGTTGGACCTTTTCAG 360

QY 1702 ValThrThrThrAlaAsnIleValGlyArgPheSerValSerIleThrGluAspIle 1721

Db 361 GTGACAACTACAGCAAAAGAGGCTGCTCTGTATCAAAAGCTGAGCAAGATC 420

QY 1722 ThrAspThrIleuSerGlnIleProValAlaIleSerProPheThrAspIleuSerGlnIle 1741

Db 421 ACTGACAAAGAAAGAAAGAGGACGATGCAATCTCTCTTTTAAATTTGAAACAAAGCT 480

QY 1742 ValIleuProAlaValIleProIleuSerGlnIleuSerGlnIleuSerGlnIleuSerIleu 1761

Db 481 GTTCTTCTGCTGTGTATACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540

QY 1762 AsnGlyProSerSerAspProGlnAlaIlePheIleuSerArgAspValAspAspGlySer 1781

Db 541 AATGGGCGGCTCTCTGACCCGAGGCGGCTTTTAAATGATGATGATGATGATGATGATG 600

QY 1782 GlySerProHisSer-ProHisGlnIleuSerSerIleuProSerGlnIleuSer 1801

Db 601 GGTAGTTCACACTCGCCCATCAGCTGAGCTCAAAAGGCTTCTTCAAGCAATCTTAAG 660

QY 1801 rGlnSerIleuSerAsnSerPheAsnSer-SerIleuSerSerAspAsnGlySerAspI 1821

Db 661 TCAAAAGCTTAGTAATTCATTTTACTCTCTTTCATGATGATGATGATGATGATGATG 720

QY 1821 leGluAspGluAspIleuSerGlnIleuSerArgIleuArgIleuArgAspIleuSerGlnI 1841

Db 721 TCGAAGTGAAGACTTAATTAAGTGTGAGCGACATGACAGATTAACCTCTCAAGAGA 780

QY 1841 leGln-AspIleuSerArgGlnIleuSerGlnIleuSerGlnIleuSerGlnIleuSerGlnI 1860

Db 781 TTCAGGAGCTGCGAGATCCCGGACAGCAAGAAATTTGATCTTTGATATCCCAACTGGG 840

QY 1860 yIleValProPro---AlaValIleIleProProAlaIleProIleuSerGlyArgArg 1879

Db 841 CAAAGTGGCCCCCTGCTGTATTAATTCCTCCCAAGCTGGCTCTCTTTCGGGGA-AAAG 899

QY 1879 GArgProThr-----LysSerIleGlySerIle 1888

Db 900 ACGACCAACCACTTAAGCAAGGAGGAGCA 933

RESULT 12

BQ213010/c 873 bp mRNA linear EST 02-MAY-2002

LOCUS BQ213010

DEFINITION AGENCOURT 7554697 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6062979

5', mRNA Sequence.

ACCESSION BQ213010

VERSION BQ213010.1 GI:20393787

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mgs.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM1335 row: f column: 04
High quality sequence stop: 648.

FEATURES
Source
Location/Qualifiers
1..873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6062979"
/issue="type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 72"
/note="Organ: Skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies.."

Alignment Scores:
Pred. No.: 3 72e-59 Length: 873
Score: 1355.50 Matches: 276
Percent Similarity: 97.89% Conservative: 2
Best Local Similarity: 97.18% Mismatches: 1
Query Match: 12.54% Indels: 5
DB: 13 Gaps: 1

ORIGIN
US-10-010-720-14 (1-2136) x BQ213010 (1-873)
306 SerGlyThr---LeuYstHrYrLeu-LysArgPhe-LysValMetLysIleLysVal 324
868 TCTGGAACCACTTAATAAGCTATCGAAGAGTTTAAAGATGATGATCAAGATTC 809
324 eu-ArgSerTrpCys-ArgGlnIleLeuYsgIleuGlnPheLeuHsThrArgThrPr 343
808 TAAAGAGCTTGGCCCGTCAGATCCTTAAAGGCTTCAGTTCTTCATATCGAATC 749
343 cProIleIleHsArgAspLeuYsgYsaSpsanIlePheIleThrGlyProThrGlyse 363
748 ACTATCATCCACCGGATCTTAAAGTACACATCTTATCACCGGCCCACTGCGCTC 689
363 rValYsgIleGlyAspLeuGlyLeuAlaThrLeuYsaArgAlaSerPheAlaLysSer 383
688 AGTCAAGATTGGAGACCTCGGTCTGGCAACCTGAAAGCGGCTCTTTGGCAAGAGT 629
383 lIleGlyThrProGluPheMetAlaProGluMetYrGluGlnYrYrAspGluSerVa 403
628 GATAGGTACCCAGAGTTATGAGCCCTGAGATGATGAGAGAAATATGATGATTC 569
403 lAspValYrYrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGluYrProTyrse 423
568 TGAAGTTATGCTTTTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGAT 509
423 rGluCysGlnSnaIaalaGlnIleYrArgValThrSerGlyValLysProAlase 443
508 GAGAGTCCAAATATGCTGGCAGATCTACCGTGGCTGACACAGTGGGTCGAAAGCCG 449

QY 443 rPheAspLysValAlaIleProGluValLysGlnIleIleGluGlyCysIleArgGlnAs 463
Db 448 TTTTACCAAAATAGCAATTCCTGAGTAGAGAAATTTTGAAGGATGATACGACAAA 389
QY 463 nLysAspGluArgYrSerIleLysAspLeuLysHsIlePhePheGlnGluLth 483
Db 388 CAAAGATGAAGATATTCATCAAAAGACCTTTTGAACCATGCTCTTCCAAAGAGAAC 329
QY 483 rGlyValArgValGluLeuAlaGluGluLysAspGlyGluYrIleAlaIleLysLeuTr 503
Db 328 AGGAGTACGGGTAAATTAAGCAAGAAAGATGATGAGACAAAATATGATTAATTTG 269
QY 503 pLeuArgIleGluAspIleLysLysLysLysLysLysLysLysLysLysLysLysLys 523
Db 268 GCTTCGATTTGAAAGATTTAAAGAAATTAAGGAAATTAAGGAAATTAAGGAAATTA 209
QY 523 uPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMetValGluSe 543
Db 208 GTTTCTTTTGAATTTAAGAGAGATGTCACAGAGATGTCACAGAGATGTCACAGAT 149
QY 543 rGlyYrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAspArgValse 563
Db 148 TGGGTATGCTGTGAGAGTATCACAGACCATGCTAAAGCTATCAAGACAGATATC 89
QY 563 rLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGluGlnGluYrLysLys 583
Db 88 ATTAAATTAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 29
QY 583 gGln 584
Db 28 AAGG 25

RESULT 13
AUI30018 825 bp mRNA linear EST 01-AUG-2002
LOCUS AUI30018 NT2RP3 Homo sapiens CDNA clone NT2RP3000108 5', mRNA
DEFINITION
ACCESSION AUI30018
VERSION AUI30018.1 GI:10990372
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 825)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
HRI human CDNA project
Unpublished (2000)
CONTACT: Takao Isogai
Genomics Laboratory
Helix Research Institute
1533-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; CDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
Source

Location/Qualifiers
1..825
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3000108"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_id="NT2RP3"
/note="Vector: pMT18SPF3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

ORIGIN

Alignment Scores:

Pred. No.: 3 38e-58 Length: 825
 Score: 1336.00 Matches: 271
 Percent Similarity: 98.91% Conservative: 1
 Best Local Similarity: 98.55% Mismatches: 3
 Query Match: 12.36% Indels: 1
 DB: 9 Gaps: 0

US-10-010-720-14 (1-2136) x AUI30018 (1-825)

QY 1642 SerThrsSerGluSerSerValLeuSerSerSerProGluSerThrLeuValysPro 1661
 DB 2 AGACCTCAGAGCTCCAGTGCATACAGAGTACGAGAGTACCTTGGAACCA 61
 QY 1662 GluProAsnGlyLeuThrLeuProGlyLysSerSerAspValProGluSerLys 1681
 DB 62 GAGCCGAAATGGCATACCAATCCCTGGTATCTCTTCAGATGTGCCAGAGTCCCA 121
 QY 1682 ThrThrAlaSerGluAlaLysSerSerProGluValProThrLysValGlyArgPheGln 1701
 DB 122 ACTACGCTCAGAGCAAGAGTACAGACTGGGAGCTTACCAAGTTGACCTTTTCA 181
 QY 1702 ValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGluAspLys 1721
 DB 182 GTGACCACTACACCAAAAGAGGCTCTTCTGATCAAAAAGTGCAGCAAGATC 241
 QY 1722 ThrAspThrLysLysGlyProValAlaSerProProPheMetLeuGluGlnAla 1741
 DB 242 ACTGACCAAAAGAGAGAGCAAGTGGCATCTCTCTTTATGATTTGAAACAAGCT 301
 QY 1742 ValLeuProAlaValIleProLysLysGluLysProGluLeuSerGluProSerHis 1761
 DB 302 GTTCTTCTCTGATGATACCAAGAGAGAGAGCTGATGATGATGATGATGATGAT 361
 QY 1762 AsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspValAspAspGlySer 1781
 DB 362 AATGGGCTCTTCTGACCCGAGAGCGCTTTTAAAGTAGGATGATGATGATGATGAT 421
 QY 1782 GlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSer 1801
 DB 422 GGTAGTCCACACTGCGCCCATCAGCTGAGCTCAAAAGAGCTCTTCTTACCGAATCTTA 481
 QY 1802 GlnSerLeuSerAsnSerPheAsnSerSerTyMetSerSerAspAsnGluSerAsp 1821
 DB 482 CAAAGCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 541
 QY 1822 GluAspGluAspLeuLysLeuGluLeuArgArgLeuArgAspLysHisLeuLysGluIle 1841
 DB 542 GAAATGTAAGACTTAAAGTTGAGCTGAGCGATGAGATGAGATGAGATGAGATGAG 601
 QY 1842 GlnAspLeuGlnSerArgGlnLysHisGlnIleGluSerLeuTyThrLysLeuGlyLys 1861
 DB 602 CAGAGCTCGACAGATCGCCAGAGCATGAAATTAATTTGTTATTCACCAATCGGCAAG 661
 QY 1862 ValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgPro 1881
 DB 662 GTGCCCCCTGCTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 721
 QY 1882 ThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerProGln 1901
 DB 722 ACTAAAGCAAAAGGACGCAATCTAGTGAAGCAATTCCTTGAGGAAATTAAGCCCCCAG 781
 QY 1902 LeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHis 1916
 DB 782 CTTTAGGTAC-CTGCTGTGTCAAAGTCAACTTCACTTTCGAC 825
 RESULT 14
 AUI31632 829 bp mRNA linear EST 01-AUG-2002
 LOCUS AUI31632 NT2RP3 Homo sapiens cDNA NT2RP3002943 5', mRNA
 DEFINITION
 sequence.
 AUI31632
 ACCESSION

VERSION AUI31632.1 GI:10991986

KEYWORDS

EST. Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 829) Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saio, K., Kawai, Y., Yamamoto, D., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and

AUTHORS

Isogai, T.

HRI human cDNA project

TITLE

Unpublished (2000)

JOURNAL

Contact: Takao Isogai

COMMENT

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

FEATURES

Location/Qualifiers

1..829

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RP3002943"

/cell_type="teratocarcinoma"

/cell_line="NT2"

/clone_id="NT2RP3"

/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

ORIGIN

Alignment Scores:

Pred. No.: 8.65e-58 Length: 829
 Score: 1338.00 Matches: 268
 Percent Similarity: 97.46% Conservative: 1
 Best Local Similarity: 97.10% Mismatches: 6
 Query Match: 12.28% Indels: 2
 DB: 9 Gaps: 0

US-10-010-720-14 (1-2136) x AUI31632 (1-829)

QY 1737 AsnGluGluGlnAlaValIleProLysLysGluLysProGluLeuSer 1756
 DB 1 GATTTGAAACAGCTGTTCTTCTGCTGATACCAAGAGAGAGAGAGAGAGAGAGAGAG 60
 QY 1757 GluProSerHisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAsp 1776
 DB 61 GAGCTTCACATCATTAATGGGCGCTCTTGACCCGAGAGCGCTTTTAAAGAGGAT 120
 QY 1777 ValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuPro 1796
 DB 121 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 1797 SerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSerSerAsp 1816
 DB 181 AGCCAGATCTTAATCAAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
 QY 1817 AsnGluSerAspIleGluAspGluAspLeuLysLeuGluLeuArgArgArgAspLys 1836
 DB 241 AATGAGTCAGATACGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 300
 QY 1837 HisLeuLysGluIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGluSerLeuTy 1856
 DB 301 CATCTCAAAAGAGATTCAGAGACTGACAGAGTGCAGAGAGCAATGAATGAATCTTTGAT 360
 QY 1857 ThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeuSerGly 1876
 DB 361 ACCAAACTGGGAGAGTCCCCCTGCTGTTATTTATTTCCCAAGCTGCTCCCTTTCAGG 420

QY	1877	ArgHrGAAGAgProThlylSerlySeryGlySerlySerSerAugsSerSerleugly	1896
Dd	421	AGMAGCACACACCCTTAAAGCAAAAGCAGCAATTTAGTCGAGAAGACTTCCTTGGGG	480
QY	1897	AsnlysserProginLeusergilyAsnuLeusergilyGinSerAlaISerValIeuhis	1916
Dd	481	AATAAAGCCCCCAGCTTTCAGGTATCCTGTGTGTGCAGAGTCAGCTTCAGTCTTGAC	540
QY	1917	ProginGlnThrIeuHisProProglyAsnIleProGiusergilyGinAsnGlnIeuleu	1936
Dd	541	CCCCAGCAGACCCCTCCACCCCTCGGCCAACATCCCAGAGTCGGGGCAGAAATCAGCTGTTA	600
QY	1937	GlnProleulyssProserProserSerAspsmIeuTyrrSerAlaPheThSerAspGly	1956
Dd	601	CAGCCCTTAAAGCATCTCCCTCCAGTGACAACTCTATTACGGCTTCCACAGTAATGGT	660
QY	1957	AlaIleSerValProserLeuseralAproGlyGlnGlyTyrrSerSerThrasnthVal	1976
Dd	661	GCCATTTCAGTACCAAGCCTTTCGTCCAGGTACAGGAACACACAGCACAAACACTGT	720
QY	1977	GlyAlathryAlanserGlnAla-AlaGlnAlaGlnProProAlaMetThrserserAr	1996
Dd	721	GGGGCAACAGTGAACAGCCCAAGCCCCGGCCCAAGCTTNAGCTTCTGNCAATGACGTCCAC-AG	779
QY	1996	glysglyThrPheThraspsIeuhislylysIeuValAspasnthrp	2011
Dd	780	GAAGGGCACAATTACAGATGACTTGCACAAAGTTGTGTGAACAATTGG	825

Mon Sep 27 08:49:58 2004

us-10-010-720-14.rst

Page 19

Search completed: September 25, 2004, 10:29:55
Job time : 7223 secs

Blank Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 00:52:38; Search time 37 seconds
(without alignments)

3005.996 Million cell updates/sec

Title: US-10-010-720-14

Sequence: 1 MSGGAABKQSGTPGSLFLSP.....NISNLSKISNPPGSLNRTT 2136

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	761.5	7.0	613	PKPA_PHYBL	Q01577 phycomyc
2	494.5	4.6	5703	MUSB_HUMAN	Q9hce4 homo sapien
3	481	4.4	1802	HKRI_YEAST	P41809 saccharomyc
4	475.5	4.4	2090	N214_HUMAN	P35658 homo sapien
5	460	4.3	1367	AMVY_YEAST	P08640 saccharomyc
6	448	4.1	5147	PCLO_HUMAN	Q9y6r0 homo sapien
7	439	4.1	2468	MABP_HUMAN	P46821 homo sapien
8	431.5	4.0	792	NEK4_MOUSE	Q921j2 mus musculu
9	429.5	4.0	3644	MINI_MOUSE	Q62504 mus musculu
10	423	3.9	5085	PCLO_RAT	Q9jke6 rattus norv
11	416	3.8	1080	NRK1_YEAST	Q9j1j9 m nucleat r
12	415.5	3.8	1080	NRK1_YEAST	Q9j1j9 m nucleat r
13	413.5	3.8	5120	PCLO_CHICK	Q9spu3 gallus gall
14	411	3.8	1239	MAK4_HUMAN	Q95819 homo sapien
15	404.5	3.7	5038	PCLO_MOUSE	Q9yqz7 mus musculu
16	398	3.7	1233	MAK4_MOUSE	P97820 mus musculu
17	396	3.7	5179	MUC2_HUMAN	Q02817 homo sapien
18	394.5	3.6	1360	TNIK_HUMAN	Q9hce4 homo sapien
19	391	3.6	5560	SPEN_DROME	Q9hce4 drosophila
20	388.5	3.6	865	CPN_DROME	Q9hce4 drosophila
21	387.5	3.6	1306	MSB2_YEAST	Q02910 drosophila
22	385.5	3.6	2464	MABP_MOUSE	P14873 mus musculu
23	385	3.6	841	NEK4_HUMAN	P51957 mus musculu
24	384.5	3.6	3664	MINI_HUMAN	Q9et58 homo sapien
25	383	3.5	2453	NCR1_MOUSE	Q60974 mus musculu
26	383	3.5	2459	MABP_RAT	P15205 rattus norv
27	382.5	3.5	1609	FIG2_YEAST	P25653 saccharomyc
28	378	3.5	939	STG2_YEAST	Q03497 saccharomyc
29	376.5	3.5	2738	PGCV_RAT	Q9erh4 rattus norv
30	375.5	3.5	1308	MAK6_MOUSE	Q9jme2 mus musculu
31	370.5	3.4	2063	NC06_HUMAN	Q14686 h nucleat r
32	365	3.4	1140	YM96_YEAST	Q14893 saccharomyc
33	363.5	3.4	4911	MLL3_HUMAN	Q8hez4 homo sapien

34	363	3.4	1332	1	MAK6_HUMAN	Q8n4c8 homo sapien
35	360	3.3	3358	1	PGCV_MOUSE	Q62059 mus musculu
36	358.5	3.3	794	1	KI11_HUMAN	Q8t0c3 homo sapien
37	357.5	3.3	1051	1	UUK1_MOUSE	Q70405 mus musculu
38	357	3.3	1196	1	HIX1_MOUSE	Q9qzr5 mus musculu
39	357	3.3	1223	1	KI25_CAEEL	Q10925 caenorhabd
40	356	3.3	1411	1	TCOF_HUMAN	Q13428 homo sapien
41	355	3.3	1198	1	HIX2_HUMAN	Q9h2x6 homo sapien
42	355	3.3	4903	1	MLX_MOUSE	Q8d7h4 mus musculu
43	354.5	3.3	3866	1	HRX_MOUSE	P55200 mus musculu
44	353.5	3.3	1258	1	NEK1_HUMAN	Q96py6 homo sapien
45	352.5	3.3	779	1	NIM1_NEUCR	P48479 neurospora

ALIGNMENTS

RESULT 1	ID	PKPA_PHYBL	STANDARD	PRT	613 AA.
AC	Q01577	PKPA_PHYBL			
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Serine/threonine protein kinase PKPA (EC 2.7.1.1-).				
OS	Phycomycetes blakesleeanus.				
OC	Eukaryota; Fungi; Zygomycetes; Mucorales; Mucoraceae;				
OC	Phycomycetes.				
OX	NCBI_TaxID=4837;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=NRRL 1555;				
RX	MEDLINE=96120859; Pubmed=8590476;				
RA	Ruiz-Perez V.L., Murillo F.J., Torres-Martinez S.;				
RT	"Pkpa", a novel Phycomycetes blakesleeanus serine/threonine protein				
RT	kinase.";				
RL	Curr. Genet. 28:309-316(1995).				
CC	-1- FUNCTION: Serine/threonine protein kinase that probably				
CC	participates as an intermediate in an intracellular system				
CC	controlling nuclear proliferation.				
CC	-1- DEVELOPMENTAL STAGE: Mainly expressed during the vegetative				
CC	growth, the level decreased when the mycelium can differentiate				
CC	and form sporangiothecae.				
CC	-1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; Z46636; CA86606.1; -				
DR	InterPro; IPR000719; Prot_kinase.				
DR	InterPro; IPR008271; Ser_thr_pkin_AS.				
DR	Pfam; PF00069; Pkinase; 1.				
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.				
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.				
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.				
KW	Transferase; Serine/threonine-protein kinase; ATP-binding.				
FT	DOMAIN 17				
FT	NP_BIND 23				
FT	FT BIND 50				
FT	ACT_SITE 140				
FT	SEQUENCE 613 AA; 69594 MW; E988308B3910A1E CXC64;				
SO	SEQUENCE				
Query Match	7.0%; Score 761.5; DB 1; Length 613;				
Best local similarity	35.6%; Pred. No. 2.7e-19;				
Matches	185; Conservative 90; Mismatches 174; Indels 71; Gaps 20;				

QY 210 KAVGNSNDGRFLKDEIEIGRSGFKTVYKGLDTETTVAVMCELDORKLT--KSERORFKE 267
 Db 6 KVIKESNGRYSKLNITVLGKAYKVYKAIDREELI-----NDNETNKKVTRQOEKRD 58
 QY 268 ---EAEMLKGIQHPITVAFYFYSWSTVKKKCIIVITELMTSGTLKTYLKKPKWM-KIKV 323
 Db 59 LGHEIDILKSVHHPNITFHDAMVNETE-----FVITELMTSGTLREYIRKLTPLPNIKI 114
 QY 324 LRSRCQILKGLQFHTPTPIIHRDLKCDNIFITGPGYSKIGLGLATLKRSFAKSV 383
 Db 115 VKRMCQILKGLAYIHGHEPPIIHRDLKCDNIFITGAHGEIKIGMGTAEMKNGK-KYTV 173
 QY 384 IGTEPFMAPEMYEEK-YDESVDVYAFGMCLEMTSEYSEYSCQNAQIYRRTSGVPA 442
 Db 174 IGTEPFMAPEMYEEKYDEKVDIYAFGMCLEMTSEYSEYSCQNAQIYRRTSGVPA 233
 QY 443 SFDEVVAPEVEIIEGCTRKQKDERYSIKDLNNAFFOEEFVGVAELAEEDGEXIAIKL 502
 Db 234 CLSRVQDELLTNTICLTP-EDERMTAQELIERFLAVEP--EVLVSKDMTKML--L 287
 QY 503 WLRJEDIKLKGKXKDNBAIEFSPDLERVDVDAQAEVSGVY-----CE----- 548
 Db 288 TLQV-----VFRG--MDKLVSVEFENADPTDADVAEMIEQVYQNCYQQLITCEINRIL 341
 QY 549 -----GDHKTMAKAIKRVSLIKRKEQOLVRE---EOKKKO--EESSLK 590
 Db 342 RDIARNOGPPDKGEDEKIVWRRENDIRSELEFRAKDLALAVERYFAEKKECELLEQNI- 400
 QY 591 QQYQSSASQGTIKQLPSASGITPTASTTSASVSTQVEPE---EPEADQHQLOLYQPSI 647
 Db 401 --IAEERCKET-IFALSOAKFQIPDLLQPOPOPOPOPOPOPOLOPOLQYLSFQS 457
 QY 648 SVLSDGTVDSGGSSVFTESRSVSOQTYSYSGQHEQASHT 687
 Db 458 TTSPPGPTSDNSTNMTLSS-LESLSKLCVSGDEQVEIT 496

RESULT 2

MUSB HUMAN STANDARD; PRT: 5703 AA.
 AC Q9HCG4; 000447; 000573; 014985; 015494; 095291; 095451; Q14881;
 AC Q99552; Q9J28;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High
 molecular weight salivary mucin MG1) (Sublingual gland mucin).
 GN MUC5B OR MUC5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE OF 1-1594 FROM N.A.
 RA Chen Y., Di Y.P., Wu R.;
 RT "Molecular cloning of the amino-terminal and 5'-flanking region of the
 human MUC5B mucin gene."
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE OF 1-1325 FROM N.A.
 RA MEDLINE=99009274; PubMed=9790959;
 RA Offner G.D., Nunes D.P., Keates A.C., Afchal N.H., Troxler R.F.;
 RT "The amino-terminal sequence of MUC5B contains conserved
 multifunctional D domains: implications for tissue-specific mucin
 functions."
 RT Biochem. Biophys. Res. Commun. 251:350-355 (1998).
 [3]
 RP SEQUENCE OF 40-1324 FROM N.A.
 RA MEDLINE=99023932; PubMed=9804771;
 RA Desseyn J.-L., Buisson M.P., Porchet N., Aubert J.-P., Laine A.;
 RT "Genomic organization of the human mucin gene MUC5B: cDNA and genomic
 sequences upstream of the large central exon."
 RT J. Biol. Chem. 273:30157-30164 (1998).

RN [4]
 RP SEQUENCE OF 1326-4895 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=97166151; PubMed=9013550;
 RA Desseyn J.-L., Guyonnet-Dupert V., Porchet N., Aubert J.-P.,
 RT Laine A.;
 RT "Human mucin gene MUC5B, the 10.7 kb large central exon encodes
 various alternate subdomains resulting in a super-repeat. Structural
 evidence for a 11p15.5 gene family."
 RT J. Biol. Chem. 272:3168-3178 (1997).
 RN [5]
 RP SEQUENCE OF 4057-4480 FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=97292540; PubMed=9147051;
 RA Nielsen P.A., Bennett E.P., Wandall H.H., Therkildsen M.H.,
 RT Hannibal J., Clausen H.;
 RT "Identification of a major human high molecular weight salivary mucin
 (MG1) as tracheobronchial mucin MUC5B."
 RT Glycobiology 7:413-419 (1997).
 RN [6]
 RP SEQUENCE OF 4721-5703 FROM N.A.
 RC TISSUE=Gall bladder;
 RX MEDLINE=97293229; PubMed=9164870;
 RA Keates A.C., Nunes D.P., Afchal N.H., Troxler R.F., Offner G.D.;
 RT "Molecular cloning of a major human gall bladder mucin: complete C-
 terminal sequence and genomic organization of MUC5B."
 RT Biochem. J. 324:295-303 (1997).
 RN [7]
 RP SEQUENCE OF 4809-5687 FROM N.A.
 RC TISSUE=Sublingual gland;
 RX MEDLINE=96125355; PubMed=8554565;
 RA Troxler R.F., Offner G.D., Zhang F., Iontcheva I., Oppenheim F.G.;
 RT "Molecular cloning of a novel high molecular weight mucin (MG1)
 from human sublingual gland."
 RT Biochem. Biophys. Res. Commun. 217:1112-1119 (1995).
 RN [8]
 RP SEQUENCE OF 4859-5703 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=97347489; PubMed=9201995;
 RA Desseyn J.-L., Aubert J.-P., Porchet N., Laine A.;
 RT "Genomic organization of the 3 region of the human MUC5B mucin."
 RT J. Biol. Chem. 272:16873-16883 (1997).
 CC -1- FUNCTION: Salivary mucin that is thought to contribute to the
 lubricating and viscoelastic properties of whole saliva.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed mainly in bronchus glands and also
 in submaxillary glands, endocervix, gall bladder, and pancreas.
 CC -1- PTM: Highly glycosylated.
 CC -1- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
 CC -1- SIMILARITY: Contains 3 VWFC domains.
 CC -1- SIMILARITY: Contains 4 C-terminal domains.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DR EMBL, AF107899; AAC36673.1; -
 DR EMBL, AF086604; AAC67545.1; -
 DR EMBL, AJ004862; CAA06167.1; -
 DR EMBL, Z72496; CAA96577.1; -
 DR EMBL, X74955; CAA52910.1; -
 DR EMBL, U63836; AAB61398.1; -
 DR EMBL, U78554; AAC51344.1; -
 DR EMBL, U78552; AAC51344.1; JOINED.
 DR EMBL, U78553; AAC51344.1; JOINED.
 DR EMBL, U78551; AAC51343.1; -
 DR EMBL, U95031; AAB65151.1; -
 DR EMBL, Y09788; CAA70926.1; -

Query Match	Best Local Similarity	20.3%;	Pred. 20.3%;	DB 1;	Length 5703;	Caps 82;
Matches 385;	Conservative 210;	Mismatches 642;	Indels 659;	Gaps 82;		
Query	596	SSASOTGICOLPSASTG-	-IPRTASTSASVSTOVEEERPADQHOOLQYQOBSISYL	650		
Db	2502	TTPTATSVTPRIPSSSLGTTWTRLISQTTTPATSTAPSSSTPETAH--	TSIVL	2552		
Query	651	SDGTVDGSGGSSVFTESS-	RVSQOQVYSGSQHEQAH-----	STGVP	691	
Db	2553	TATATTGATGSVATPSTPGTAHTTVPTTTTGTATPSSPGTALPPVWISTTTP		2612		
Query	632	GHIPSTVQAQOPHGV-----	PPSSVNOG-----	OSQGPS-	SSLTGSSS	733
Db	2613	TTRGTTVPSSIPGTHTAIVLTTTITVAATGSMATPSSQTQSGRPPLITATTITTA		2672		
Query	724	QPIHQOQOQOQIQAAPPOQTVOYSLQSTSTSEATTAQPVSOPOAPOVLPVVSAGKOST		793		
Db	2673	GSTINPSSPTGRPRIPPLVLT--	ATTPATSSIVTPSSALGTHTRPPV--PNTTA--	TT	2726	
Query	794	QGVSOVAPAEPAVAV-----	AQPAQTQPTTLASSVDA-----	HSVVASG--		832
Db	2727	HGRS-LSPSSPHVTAMTATSGLTGTHITESTGSHUPATTTGTQHSPTASSPH		2785		
Query	833	-MSDGENVPSSSGHGEKRT--	TRHVKYSVRSR-----	HEKTSRPKLRILNVS--		880
Db	2786	PSRRTTESPPSGITTPHTTATSKTTATAPSKTRISTLLPSSPAPLITTVVTMGCER		2845		
Query	881	-----	NKGD-----			RV
Db	2846	QCAMSHMLDYSPYMGPSGGPDTYSNIRAGAVCEQPLGLECRAOAGPVPLRELGOV		2905		
Query	887	VECOLE-----	THRRKV-----	TEKFDLD-----	GNPELAIITMNNDFILIERESF	931
Db	2906	VECSLDPLFGLVCRNRQVQKFKMCFMYELIRVPCNNGHCSTPAT--		2949		
Query	932	VDOVRRIIEKADMEISDVSEPEEDQG-----	LESLOKGDYGFSSGOKLEGEFKOP	IP		986
Db	2950	-----	SSATPSSPTGTTMLIBQTTAATTATGSGTALPSSPTGPAP			2992
Query	987	ASS-MEQIGIFPSSILQVHSGAGRFIV-----	SPVPSRL-----			1022
Db	2993	PPKVLISQATTPATSSSSSRRLATTLVLSTAKSTATSPFIPSSITGTTGTS		3052		
Query	1023	-----	RESKVPSEITDTVAASIAQSPKAMLSHASSLSLOQA-----	FSELR		1066
Db	3053	QNRPHPMATMSTIHSSPETHSTVLTATKATTRATSSGSTSSSPGTTMLITELTT		3112		
Query	1067	AQMTGEG--PNTAPNNSHGPTFPVAVPEPLSIAGVPTTAATAVAPATSSPPNDISTSV		1124		
Db	3113	AAITTAALPHGTPS--STPGTTMLILBESSTIAVTVVPGSTATAS--		3155		

FT CAROHYD 1400 1400 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 1802 AA; 188890 MW; E344CA6469785A24 CRC64;
 Query Match 4.4%; Score 481; DB 1; Length 1802;
 Best Local Similarity 22.6%; Pred. No. 2.3e-09;
 Matches 333; Conservative 210; Mismatches 603; Indels 328; Gaps 64;

QY 596 SASAQGTGKQPLPASTGIPITASTASVSYQVEPEPEADQHQOL--QYQCSISVLSQD- 652
 DB 29 STSYNNGIESTPSYST--SAISTGSSNKENAITSSSETTTMAGQYGEGSTTIND 83
 QY 653 -CTVDSGGSSVFTESRVSSQOTVSYGSGHQEAHSTGVPGHPIPTVQAQSGPHGV---- 707
 DB 84 QETGHSQYIYVTTTGTQSDTMS-SYKSTELIATPESSI---VPTLQGSDESQISQTL 139
 QY 708 -YPPSSVAQGGQGGQPS-SSSLTGVASSQPIHQPOQGGIQQTAPQQTQVQVLSQSTPS 765
 DB 140 SHNPKSVASDSDPTTSSSSSSSVIISTSDSAVPREISPIITTDQISKEEGTLAQTSI 199
 QY 766 SEATTAQPVQPOAPQVLPQVSGKOSTQGVQVAPPAVVAQPOATQPTTLASV--- 822
 DB 200 SEET-----RIAGVTRVSO-----ISITLASTIDGSSSTO-TDPSNTVSE 243
 QY 823 DSAHSDVAGMSDGNENVPSSSGRHEGRTTKRHRYKSVSRHEKTSRPKLILNVSNK 882
 DB 244 NSVEEYAMSKQLSSESYSSSTVYSG-----GESTADKTSSP--ITSFSS 289
 QY 883 GDRVECOLETHNRKMYTEKFDLDGDNPEIATIMVNDPILAESEFPVQREIIEKA 942
 DB 290 YSQTTSTETSESRVAV-----GVSRPSITQTTSI--- 320
 QY 943 DEMLSDEVYVPEPGDQLESLOGKDYGFSGSQKLEGEFKQPIPASMPQOI-----GI 996
 DB 321 DSEFMSSEVELSTY-----YDLSAGNVPDQELIVDRPAISTTAEISSEASQV 367
 QY 997 PTSSLTQOVHS-AGRRITVPVPESTRLEKVPFSEITDVAASTAQSGMNLSHSASLT 1055
 DB 368 SRESNTAVASISITNTIVSSASDT-----VISTSTNTVPVSSVMTVHAHTSSSTYL 421
 QY 1056 SLOQAFSELRRQMTGEPNTAPPNFSHTGTFPVVPPFLS--SIAGVPTTAATAPVA 1112
 DB 422 S-----SSLYSSPFLSASVSHFG-----VAPFPAYISFSSVPVAVSST---Y 462
 QY 1113 TSSPPNDISTSVIOSEVTPTEBGIAGVANTGVVTSGGGPIPVSSPVLSSVSSITI 1172
 DB 463 TSSP-----SASVVP-----SAYAS-----PSVPA---VSTITSSPSA 496
 QY 1173 PAVVIS--TTPSLQVPTSTSEIVSS--TALYPSVTVASASAGSTATPGKPPA 1227
 DB 497 PAIISSTYSSPS--APVAVSYTSSPSAPAISSITYTSSPSAPVAVSSTYSSPSAPA 554
 QY 1228 VVSOQAAAGSTTVGATLTSVSTTSPSTASQSLQSSSTSTPLAETVTVVSAASLQKTS 1287
 DB 555 AISTYTTSSPS--APVAVSYTSSPSAPAISSITYTSSPSAPVAVSSTYSSPSAPA 612
 QY 1288 HSGTTG-----LASLSAPSSSSSPGAVSYIQPGHLPVSVIAPSTPIIPQAAP 1342
 DB 613 SSTYSSPSAPVAVSSTYSSPSAPAISSITYTSSPS--VFAVSYTSSPSAPAISS 670
 QY 1343 TSTPLPQVPSIPPLVQPV--ANVPAVQOTLIHSQOPALLPQPHHCPEVDS--- 1394
 DB 671 TYT-----SSPVPVAVSSTYSSPSAPAISSITYTSSPSAPVAVSSTYSSPSAPA 726
 QY 1395 --DTQPKAPGIDIKTLEKRLSFLSEHSSGCAHASVLETSVYESTVPGIPTTVA 1452
 DB 727 TYTSSPSAP-----VAVSYTSSPSAPAISS-----STYSS--PSAPVA 765
 QY 1453 PSKLTISTTSTCL---PPTNPLGTVALPVT-PVVTGGQVSTPVSTTSGVKGPGTAPSK 1507
 DB 766 VSYTSSPSAPVAVSSTYSSPSAPVAVSSTYSSPSAPAISSITYTSSPS--SASVMSSTSSS 822
 QY 1508 PLTKAPVLPVGTETLPAGLTPEEQLPPEFGPSLTQSQOPLDLDQAQLRTTSPKXITVTS 1567

DB 823 P--YDIYVSSSSASRSSTATYEFSPSPSTSLPTS-STYTFSSAVAFBESSERYSTTS 878
 QY 1568 AVGPVSM-AAPTATTEACTOPQKGVSYQVEGPVAT-----SSGAGVFMEKRFQVSV 1618
 DB 879 TIAPRQHISTLSRITDFLQTSMAIQSVSQOISTSTLNDIHSALSVF----- 929
 QY 1619 AADQAKGKKNKSEDAKSVHEESSISESSVSSSSPESTLVKPEPNGITIPGISDVPS 1678
 DB 930 -----NPS-----ASNIVETSLIIST-QASITSPK-NSAKISLQSQSSS 969
 QY 1679 AHK--TTASEAKSDPTQKGRFQVTTTANKVGRFVSQKTEDKTIDTKKEGPVAPPPM 1736
 DB 970 TKNPYDTANKEITISGRSTVSNFYTSSAKPKNDKESATPTELT----- 1015
 QY 1737 DLEQAVLPAVTPKKEKPELSEPSH--LNGPSDPEAFLSDVDGSGSP---HSPHQ 1789
 DB 1016 -----TSSSHAYSLIFSHNSVYGLSHNPFDSKSNATSGFSSSS 1058
 QY 1790 LSGKSLPSQNL--SGSLN-----SPNSVYSSDNEEDIEEDDKLELRRLRDKHLKEI 1841
 DB 1059 ISSIKLSKETIPASKSVSNTQERITSFTSLRAN--SQSEKSEGNVSGSLQSSHISN 1115
 QY 1842 QLOSRQKEIETSLYTKLGKVPVAVIIPPAELSGRRRPTKSKGSSRSSSLGNKSPQ 1901
 DB 1116 PSLSTVTKVDSKSLSKVKTGGE-----NGEFTLTTTKQYKSSSTSGSYRS 1166
 QY 1902 LSGNLSSGASAVLHPQTLHPGNIPESGQNLQPLKPSPSQNLVYAFISDGLISVP 1961
 DB 1167 FT-KISIGATTAVQCASTNSVETAPALS---TYPTTPYP-SPNSVAMLPT--AIIYE 1218
 QY 1962 SUSAPGQGSTTNTVATVNSQAQQAQPPAMSS 1995
 DB 1219 S-SETGPTTASFN---PSITGLPNAIPEAVAVS 1248

RESULT 4
 N214 HUMAN
 ID N214 HUMAN STANDARD; PRT; 2090 AA.
 AC P3658;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Nuclear pore complex protein Nup214 (Nucleoporin Nup214) (214 kDa
 DE Nucleoporin) (CAN protein).
 GN NUP214 OR CAN OR CAIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=92195315; PubMed=1549122;
 RA Von Lindern M., Formero M., Van Baal S., Jaegle M., De Wit T.,
 RA Bulj A., Grosveld G.,
 RT "The translocation (6;9), associated with a specific subtype of acute
 RT myeloid leukemia, results in the fusion of two genes, deK and can,"
 RT and the expression of a chimeric, leukemia-specific deK-can mRNA,"
 RL Mol. Cell. Biol. 12:1687-1697 (1992).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=94151361; PubMed=8108440;
 RA Kremer D., Wozniak R.W., Blobel G., Radu A.,
 RT "The human CAN protein, a putative oncogene product associated with
 RT myeloid leukemogenesis, is a nuclear pore complex protein that faces
 RT the cytoplasm,"
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1519-1523 (1994).
 CC - FUNCTION: May serve as a docking site in the receptor-mediated
 CC import of substrates across the nuclear pore complex.
 CC - SUBUNIT: Homodimer. Interacts with DDX19 and NUP88.
 CC - SUBCELLULAR LOCATION: Nuclear pore complex. Cytoplasmic filaments.
 CC - TISSUE SPECIFICITY: Expressed in thymus, spleen, bone marrow,
 CC kidney, brain and testis, but hardly in all other tissues or in

CC whole embryos during development.
 CC -1- DOMAIN: Contains F-G repeats.
 CC -1- P1M: Probably glycosylated as it reacts with wheat germ agglutinin (WGA).
 CC -1- DISEASE: Implicated in a subset of acute myeloid leukemia (acute nonlymphocytic leukemia) (AML) carrying a chromosomal translocation t(6;9)(p23;q34) that results in the formation of a DER-CAN fusion gene.
 CC -1- DISEASE: Involved in some cases of acute undifferentiated leukemia (AUL) through a chromosomal translocation t(6;9)(q21;q34.1) that involves NUP214/CAN and SET.
 CC -1- DATABASE: NAME=Atlas Genet. Cytoenet. Oncol. Haematol.;
 CC -1- WWW="http://www.infobiogen.fr/services/chromancer/Genes/CAN.html".
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X64228; CAA45535.1; -
 CC PIR: S26058; S26058.
 CC Genew: HGNC:8064; NUP214.
 CC MIM: 114350; -
 CC GO: GO:0005643; C:nuclear pore; TAS.
 CC GO: GO:0005215; F:transporter activity; TAS.
 CC InterPro: IPR004325; Nucleoporin_Fg.
 CC InterPro: IPR01680; WD40.
 CC Pfam: PF00393; Nucleoporin_Fg; 22.
 CC SMART: SMO0320; WD40; 2.
 CC KW Nuclear protein; Transport; Proto-oncogene; Chromosomal translocation; Repeat; Glycoprotein.
 CC FT DOMAIN 481 2076 11 X 5 AA APPROXIMATE REPEATS.
 CC FT DOMAIN 1409 2084 18 X 4 AA APPROXIMATE REPEATS.
 CC FT DOMAIN 1427 2085 11 X 3 AA APPROXIMATE REPEATS.
 CC FT DOMAIN 1213 2090 PRO/SER/THR-RICH.
 CC FT DOMAIN 740 768 LEUCINE-ZIPPER 1.
 CC FT DOMAIN 861 882 LEUCINE-ZIPPER 2.
 CC FT SITE 812 813 BREAKPOINT.
 CC SEQUENCE 2090 AA; 213766 MW; 6DBE767FDD857F8F CRC64;
 QY Query Match 4.4%; Score 475.5; DB 1; Length 2090;
 Db Best Local Similarity 20.7%; Pred. No. 4.2e-09;
 Matches 332; Conservative 238; Mismatches 740; Indels 525; Gaps 79;
 QY 545 YVCEGHHKTKAKIKORVSLIKRKRQ-----RQVAREOEKKKQOEESLKKQVEGS 596
 Db 309 YIEEMVLVLAASASTEVSTILARQSDQINWESWLBEDSSRAELPVYDKSDPSLPMGVVD 368
 QY 597 SASQTGI-----KOLPSA-----STGIPTASTTSASVTOYE--- 628
 Db 369 YINQVETITSEKTLPPAPVLMSTDVLCPFYMINQNGVSLIKTERISLSEEROP 428
 QY 629 -----PEEPBADHOQ-----LQYQPSISVLSGTVDGQSSSVFT--ESR 668
 Db 429 KSPGSTPTTTSQAPQKXDASAAAPASLPPSSPAPIATFSLIPAGAPVTFSGSSS 488
 QY 669 VSSQQTV-----SYSGHQASHSTGVPHISTVQAGQPHGVPPSSVA---QGSGQ 719
 Db 489 LKSSATVTGEPSPSYSSGSD--SSKAAPRPGPST-----FSFVPPSASLAPTPAASP 538
 QY 720 GQPSSSSLTGVSS-----SQPIHQOQOQ-GIQOTAPQ-QTQVSLSCSTSTSEAT 770
 Db 539 VAPSAASFSGSGFKPTESTPVPSVSNPNTAMSSFPSPSAYKAMISEKFTA--AAT 596
 QY 771 AQPVQPC-AEPUVLQVSAQKQSTOG-VSQVAP--AEPUVAQOPATOPTTLASSVDSAH 826
 Db 597 STPVSSSGAPPMSPFSSASKEPASHPTPLAPPSVPLKSSVLPSPSGRAGQSS 656
 QY 827 SDVASGMDGNEVNSSSGRHEGRTTKKHVRSVRSRHE-KTIRPKRLINLVNSNKDR 885

Db 657 SPVSEVQKSPRIITPPA--KFGSPQAKSLQPAVAEKQHQWCKSDPVM-----AGIGHE 709
 QY 886 VVECOLETHNRKMTVTEKEDLDGDNPEEATIMVND-----FYIAER----- 928
 Db 710 IAHQKELEELKARISKACQFQVIGSEKKMLRTESDHLITLLEIKETTESLHGDISSLK 769
 QY 929 -----ESF--VQVREIIEKADQM-----LSEDVSEPEGDQLESIQGKDYDFSGSQK- 976
 Db 770 TTLLEGFAGVEAEQRNRNDSGYLHLVPRPLDPKSEAOIQEIRLHQYKFAVQVUN 829
 QY 977 --LEGFKQPIIPASSMPQOQIIP-----TSLTQVHSAGRF--TVSEVPSRLRE 1024
 Db 830 DVLDEMDOHLEQKKQKHILVPERETLFTLANRREIINOQRKRLNLVLSLOQLRLYK 889
 QY 1025 -----SKVPSE-----ITDVAASQSPGNLSHSASLSLQOAFSELR 1065
 Db 890 QTSLSMSIAVPSQSSIHSPDLESICNMLKTTIE-----SHTKSLPVPAKLSFMK 943
 QY 1066 RAQM-----TEGP--NTAPNFSHTGPTTPVVPPLS-----SIAGVPTTAAT-- 1107
 Db 944 QOQLRNLAKRKTPPVASTAPASLSRSA-----FISQRYEDLDEVSTSSVSGLE 995
 QY 1108 -----APVPATSSPNDISTSVI-----QSEVTVPTRETAG- 1139
 Db 996 SEDARTCKDDEAVYQAPRHAPVVRTPS-----IOPSLPHAPPAKSHLVHGSSPGVWGT 1051
 QY 1140 -VATSTG-VVTSG-----GLPIP--PV-----SESPVLSV 1166
 Db 1052 SVATISAKKIIPQAGADSTMLATKTVKKGAPSPSHISAPQQLAALARRQMASQAPAVNTL 1111
 QY 1167 VSSI--TIPAVVISITTSPELOVP--TSTSEIVVSTLAYS--VVSATSSAGSGSTAT 1220
 Db 1112 TBSLTKNVPQVAVVVOELKNNPATPSTAMGSSVYSTKTHPVLTPLVAANQAKGSLINS 1171
 QY 1221 PGKPPPAVVSQQAAGSTTVGATLTSVTTSPSTASQSLQSSSTSTTLAETVVA 1280
 Db 1172 LKSGGTPPASGQSLSDSKASGTAKIETAVTSTBSAGQSKPSPSPSGTGNFGIITPT 1231
 QY 1281 HSLDKTSHSSTTGLAFSLSPASSSSPPGAGVSSYISQPGIHLVLPVSLASPILPOA 1340
 Db 1232 PSENFMAGQANPSTKSSQPDFAFSSGGGSKPEYEAIPESPPSGITSAINTTPGEEAAS 1291
 QY 1341 GPTSTPLPQVPSIPPLVQPVANVPA--VOQTLHSGOPALPENOHTHCEVDSPTOP- 1398
 Db 1292 --SSRVAASGTALSTSSKLETPPKLGEHLFPSSLAGETLGSFGLRGVQADDSKTP 1349
 QY 1399 -KAPGIDDIKLEEKLSLSEHSSSG--AQHASVLSLTSVITSTVPGIPTAVAP 1453
 Db 1350 NKASSTSLSTQPTKTSQVPSGFNFTAPVVLGKHT-----BPPVTSATTTSVAP 1399
 QY 1454 SKLLISTSTICLPNTLPIGTVALPVTPVTPGOVS--TPVST-----TTSGVK 1500
 Db 1400 -PAATSTSTAV-----FG--SLPVTASGSGVLSFGSTLSAGKTSFSGSQQINSTV 1450
 QY 1501 PGTAPEKPLTKA-----PVLVGTGLPAGTLSEQLPPEPBGSS--LTSGQPLEDL 1550
 Db 1451 PPSAP--PPTTATPLPISFPLSFGSLSSATTPS--LPMAGSTEBEATSSALPEKPG 1506
 QY 1551 DAQLRRLTSPEXITVTSVAVGPMAPFAITEAGT--OQKGVSYKKGVLATSSG 1605
 Db 1507 DSE-----VSASAASILEQOQSAQLPOAPQTSIDSYKKEPVL----- 1543
 QY 1606 AGVFMMKGFQVSVADGQKBEKKNSEDAKSVHPESTSESSVSLSSSPES--TLVKPEBN 1664
 Db 1544 -----AQPVVSNGTAAASTSIALVAEAETPATVGPDAATEAVPPAS 1586
 QY 1665 GITTPG-----ISSPVESAHTTASSEKSDP-----GQPTVGFQVTTANKV-- 1709
 Db 1587 SFSVQGVTAVALAASISAGPAVETSTPLASSTSIYAPGSAEAAAFGTVTSGSSVFA 1646
 QY 1710 --GFSVSKTEKXITDVKKEGPVVASPPFMDLEQAVLPVITKK-----EKPELSEP 1758
 Db 1647 QPRAASSANQULNTNNTATAPASATPVFGVAASTAPSLFGQQTGSTASTAAATPQVSS- 1705

QY 1759 SHLNGSSDPEAFSLRDVDDGSGSPHSHQSLPSQNLQSLNSFNSSYMSDNE 1818
 Db 1706 SFPSSBAFTTAPVPGGQTTFOQASV-----FCQASSASAVFSFSGPFG 1750
 QY 1819 SDIEDEDLRLRLRDLKELKEIODLSHQKHEISLYTKLGKVPVAVIIPPAALSGRR 1878
 Db 1751 SSV-----PAFQGPASS----- 1762
 QY 1879 RRPYSKGS-----KSSRSSLGKNSPQLSGNLGSGASVHLHPQTLHPGNIPES 1930
 Db 1763 -PPTSTSGSVFGAASSTSSSSSFSGQSSPNTGGILFGQS-----NAPAF 1806
 QY 1931 GGNOLLPLKPPSSSNLYSAFTSDGALVPSLS-----APGQSTSTNTY----- 1976
 Db 1807 GQ-----SPFGQGSVFGGTSAAATTATTSFSCQASGFSSTGVSFGQAAT 1857
 QY 1977 -GATVNSQAQAQAPPMATSSRKGTFTDHLKLVDMNARDAMLGRRGSKGHNNYEGPM 2035
 Db 1858 GGIVFGQSSSSSGSVFGSGNTGRGGGFGSGLGKXSGQAANKNPSSASGFG----- 1910
 QY 2036 ARKFSAPGQLCTSMSTNLGSGAPISAAATSLGHFTKSMCPPOGYGFPATPFQAQMSGTG 2095
 Db 1911 -----GSTATSTNTSLFNGN-----SGAKTFGGFASSSFGEGK-----PTGTSSGGG 1952
 QY 2096 GPAPQPLGQFQPVGTASLQNFNISNLQKSISSNPPG 2130
 Db 1953 SVASQGFSSSPKKTGGFGAAPVPGSHPPTGSSPG 1987

RESULT 5
 AMTH YEAST STANDARD; PRT: 1367 AA.
 ID AMTH YEAST STANDARD; PRT: 1367 AA.
 AC P08640; P08068;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glucoamylase SI/S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).
 GN STAL OR STAZ OR MAUS OR YIK019C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=9731266; PubMed=9169870;
 RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D., Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N., Harris D.E., Horenell T., Hunt S., Jagsels K., Jones M., Lye G., Moulé S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N., Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
 RL Nature 387:84-87(1997).
 RN [2]
 RP SEQUENCE OF 1-242 AND 762-1331 FROM N.A.
 RC MEDLINE=87194600; PubMed=3106330;
 RA Yamashita I., Nakamura M., Fukui S.;
 RA "Gene fusion is a possible mechanism underlying the evolution of STAL.";
 RL J. Bacteriol. 169:2142-2149(1987).
 RN [3]
 RP SEQUENCE OF 1-31 FROM N.A.
 RC STRAIN=SPX101-1C;
 RX MEDLINE=89031230; PubMed=3141213;
 RA Pardo J.M., Ianez E., Zalacain M., Clares M.G., Jimenez A.;
 RA "Similar short elements in the 5' regions of the STAZ and SGA genes from Saccharomyces cerevisiae.";
 RL FEBS Lett. 239:179-184(1988).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively from non-reducing ends of the chains with release of beta-D-glucose.

CC -1- SIMILARITY: TO S. POMBE SPBC215.13.
 CC -1- SIMILARITY: SOME, TO S. POMBE SPCC285.13C.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z38061; CAA86176.1; -
 DR EMBL: M16164; AAA35014.1; -
 DR EMBL: M16165; AAA35015.1; -
 DR EMBL: X13857; CAA32069.1; -
 DR PIR: S48478; S48478.
 DR GeneBank: 139731; -
 DR SGD: S0001458; MUC1.
 DR GO: GO:0005886; C:plasma membrane, IDA.
 DR GO: GO:0030447; P:filamentous growth, IDA.
 DR GO: GO:0007125; P:invasive growth, IMP.
 DR GO: GO:0007124; P:pseudohyphal growth, IMP.
 KW Hydrolyase; Glycosidase; Polysaccharide degradation; Glycoprotein;
 KW Signal; Multigene family.
 FT SIGNAL 1 21
 FT CHAIN 22 1367
 FT DOMAIN 210 1367
 FT CARBOHYD 817 817
 FT CARBOHYD 874 874
 FT SEQUENCE 1367 AA; 136110 MW; 91C00E2DBD61AA9D CRC64;

Query Match

4.3%; Score 460; DB 1; Length 1367;

Best Local Similarity 22.4%; Pred. No. 8; 8e-09;

Matches 302; Conservative 164; Mismatches 540; Indels 342; Gaps 54;

QY 585 EESLSAQVHQS--ASQCIRKLPASGIRPASTTSAS-----VSTQVPEEPEAD 635
 Db 229 ESSSTTSSTSSSTSSSTSSSTSSSTTAPATPTTSTCKEPTPTTSTCKEPTPT 288
 QY 636 QHQQLQVQPSISVLSLDTGTVDSQGSVPFTSSVQQTVSYGSHQAHSTGTVPCHP 695
 Db 289 HHD-----TTPCTKKTKTSTCT-----KXTTPVPPTSS 319
 QY 696 STVQAGQPHGVYPSSVAQGGQGGQSSSLGVSSSQPTQHPQOQGGIQAAPPGQTV 755
 Db 320 STSSSAP--VPTSSSTSSSAPVTSSTT-ESSSAPVPTSSSTSSSAP----- 370
 QY 756 QYLSQTSSTSEATTAQVQAPQVLPQVSAQKSTQGVSO--VAAPVVAQAQQA 812
 Db 371 -VTSSTSSSAPVTSSTSSSAPVPTSSSTSSSAPVTSSTSSSAPVTSSTSS 429
 QY 813 TQPTLASVDSASDVASGMSDGN--VPSGSGRHEGRTTKRHYKSKVSRSHKXT 869
 Db 430 SSAPVTSSTSSSAPVTSSTSSSAPVPTSS-----T-----ESS 469
 QY 870 SRPRLRIANVSNKGRVVEQLTHNRKMTTFKEDLDGNPELATIMANNDFILALERE 929
 Db 470 SAP-----VTSSTSSSAPVPT-----FSSST-----ESS 496
 QY 930 SFVQVKEIEIKADENLSEDEVSEPEGGQGLSLQKQDVGFSGSQKLEGFKQPIASS 989
 Db 497 S-----APVTSSTSSSAPVPT-----SSSTSSSAPVPTSSS 533
 QY 990 MPQOIGIPTSLTQVHSAGRRFVSPVESRLRESKVP--SEITD-----TVAAT 1040
 Db 534 TTSSSAPVTSSTSSSAPVPTSSSTSSSAPVPTSSSTSSSAPVPTSSSST 588
 QY 1041 AQSPEMNISHSASSISLQAFSELRRAQMTGCPNTPAPNFHT--GPTTPVVPPL-----S 1095
 Db 589 TESSSAPVPTSSSTSSSAPVPTSSSTSSSAPVTSSTSSSAPVPTSSSSTES 648
 QY 1096 SIAGVP-----TVAATAPVATSPPDISTSVQSVTVPTGEGIGAVTSGVTVTS 1150


```

Db 649 SSAPVPTSSSTESSAPVPTSSSTESSAPVTSST---TESSAPVTSST--TESS 703
QY 1151 GLPIPVESPVLSVSSVSTIPAVVSTSTSPSQVPT-----SSEIVSVSTALYP 1203
Db 704 SAPVPTSSSTES---SSAPVPTSSSTESSAPVPTSSSTESSAPVTSSTESS 760
QY 1204 SVTVSATASAGGATATGPKPRAVVSQAAGSTVGAULTSVSTTSSFPSTASQLSIOL 1263
Db 761 SAPVPTSSSTESSAPVPTSSSTESSAPVPTSSSTESSAPVPTSSSTESSNTS 820
QY 1264 SSSTPTPLAEVTVVVAHSIDKTHSHSTTGAFSLASASSSSSPGAVSSYISQPGGLHP 1323
Db 821 SAPSTPTSSST---ESSVPVPTSSSTTE---SSAPVSSSTESSVAP----- 865
QY 1324 LVIPSVIATPILPQAPGTSTPLLPQVPSIP--PLVQPVAN--VPAVQ--QLIHSGQPPA 1379
Db 866 --VPTSSSSNTSSA--PSSIPSTSTESSTGTVTPSSSKIPGSGTETSVSTETT 921
QY 1380 LLPQPHHCEVDSIDTOPKAPGIDDKTLEKRLSIFSEHSSGQAQHASVLETSVIE 1439
Db 922 IVPTK-----TTTSVTPSTTTITTTVCSGTNSAGETSGCSPKTV----- 963
QY 1440 STVPGIPTTAVAPSKLSTSTSCLPPTNLPLGTVAPLPTVPTVQGVST--PVSTT- 1495
Db 964 TTVPTTTTTSVTTSTTTTIVCSGTN---SAGETTSQSPKTTTIVPCSTSPS 1018
QY 1496 -TSGVKGTAPSKPPLTKAPVLPGVTELPAGTLPSQLP-----PQPSLQSQPLE 1548
Db 1019 ETASESTTSPTTIVTVSTTVTTSTSTKGGETTTFVTKNIPPTVTL- 1071
QY 1549 DLDAQRLTSPKRTVTSVAVGPMAPPAIT---EAGT---QPKGVQVKEGPVL 1600
Db 1072 -----TTPVTVTFNPTPTTTTIVCSGTNSAGETSGCSPKTVTTV 1118
QY 1601 ATSSGAVFMKGFQVAVAA-----DGAQKGRKKS--EDASVHPESTSGSV 1648
Db 1119 PCGTGCEYTTATLTAVTTVTTVTSSTGNSAGKTTGYTKTSV---PTTVTT 1174
QY 1649 LSSSSPESTLVKEPENGITIPGISDVPESAHKTASAKS----- 1689
Db 1175 LAPSAPVTPATNAVPTTTTTESSAATNAGETTSVCSATIVSSASAGNTAPSATPV 1234
QY 1690 DTGQPKVGRFQYTTANKGRFSVKTEDKIT-----DTKKEGPVAPSPFMDLEO 1740
Db 1235 TTAIPT-----TITTESVGTNSAGETTTGYTKSIPTTYITLLPGSAGAKNYETVAN 1289
QY 1741 AVLPAYIPKKEK-----PELSEPHNLGSPSDPEAAFLSDVDDGSGSPHSPHQ 1789
Db 1290 ANTPISIKTTSQLATTASASVAPVTSBS-LTGP-----LOSAGS-----A 1331
QY 1790 LSSKSLPSQNTLSQSLNSFNSSYSMSDN 1817
Db 1332 VATYSVP-----SISSTYQGAN 1349

```

RESULT 6

```

PCLO HUMAN          STANDARD; PRT; 5147 AA.
ID PCLO_HUMAN          Q9Y6V0; O43373; O60305; Q9BYC8; Q9YIV2; Q9Y6I9;
AC Q9Y6V0; O43373; O60305; Q9BYC8; Q9YIV2; Q9Y6I9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Piccolo protein (Aczonin) (fragments).
GN PCLO OR ACZ OR KIAA0559.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE OF 1-759 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99439764; PubMed=10508862;
RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,

```

```

RA Kilimann M.W.;
RT "Aczonin, a 550-kd putative scaffolding protein of presynaptic active
RT zones, shares homology regions with rim and bassoon and binds
RT profilin."
RL J. Cell Biol. 147:151-162(1999).
RN [2]
RP SEQUENCE OF 552-4404 FROM N.A.
RA Kremer U., Wollam C., Woldmann P., McGrane B.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 3619-5147 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RL code for large proteins in vitro."
RN DNA Res. 5:31-38(1998).
RN [4]
RP SEQUENCE OF 4405-4439 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Stapleton M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggiano N.A., Peters G.U., Abramson R.D., Mulhaly S.J.,
RA Bork S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Botterfield J.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 4405-5147 FROM N.A.
RA Kallio J., Elliott G.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May act as a scaffolding protein involved in the
CC organization of synaptic active zones and in synaptic vesicle
CC trafficking (By similarity).
CC -!- SUBUNIT: Interacts with Rabac1, Pral, RIMS2 and profilin (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
CC synaptic junctions (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing. Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=Q9Y6V0-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9Y6V0-2; Sequence=VSP_003923, VSP_003924, VSP_003925,
CC VSP_003926, VSP_003927;
CC Note=No experimental confirmation available;
CC -!- DOMAIN: C2 domain 1 is involved in binding calcium and
CC phospholipids. Calcium binds with low affinity but with high
CC specificity and induces a large conformational change.
CC -!- SIMILARITY: Contains 2 C2 domains.
CC -!- SIMILARITY: Contains 1 PDZ/DRH domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; Y19188; CAB60727.1; -
 DR EMBL; AC004903; AAD20936.1; -
 DR EMBL; AC004886; AAD21789.1; -
 DR EMBL; AB011131; BAA25485.1; -
 DR EMBL; BC001304; AAH01304.1; -
 DR EMBL; AC004082; AAB97937.1; -
 DR PIR; T00634; T00634.
 DR HSSP; P04410; 1A25.
 DR Genew; HGNC:13406; PCLO.
 DR MIM; 604918; -
 DR GO; GO:0005856; C:cytoskeleton; NAS.
 DR GO; GO:0045202; C:synaptic junction; ISS.
 DR GO; GO:0005509; F:calcium ion binding; ISS.
 DR GO; GO:0005544; F:calcium-dependent phospholipid binding; ISS.
 DR GO; GO:0005522; F:profilin binding; ISS.
 DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
 DR GO; GO:0016080; P:synaptic vesicle targeting; ISS.
 DR InterPro; IPR000008; C2.
 DR PRINTS; PR00360; C2DOMAIN.
 DR PRINTS; PR00399; SYNAPTOTAGMIN.
 DR SMART; SM00239; C2_2.
 DR PROSITE; PS00499; C2_DOMAIN_1; 1.
 DR PROSITE; PS50004; C2_DOMAIN_2; 2.
 DR Calcium/phospholipid-binding; Zinc; Metal-binding; Zinc-finger;
 KW Repeat; Alternative splicing.
 FT NON TER 1
 FT DOMAIN 400 465
 FT 10 X 10 AA TANDEM APPROXIMATE REPEATS OF
 P-A-K-P-O-P-Q-O-P-X.
 C4-TYPE (POTENTIAL).
 C4-TYPE (POTENTIAL).
 POLY-PRO.
 PDZ.
 C2 DOMAIN 1.
 C2 DOMAIN 2.
 S -> SGNGLGIRIVGKEIPIGHSGETIGAYIAKILPGSAAE
 QTGKLMG (in isoform 2).
 K -> KPDDGKVVSHITGEBIQ (in isoform 2).
 FTRID=VSP 003923.
 FTRID=VSP 003924.
 G -> GQVAVYQNAS (in isoform 2).
 FTRID=VSP 003925.
 TAHS -> SKRRK (in isoform 2).
 FTRID=VSP 003926.
 Missing (in isoform 2).
 FTRID=VSP 003927.
 FT SEQUENCE 5147 AA; 563537 MW; CDS084390498CD3C CRG64;
 Query Match 4.1%; Score 448; DB 1; Length 5147;
 Best Local Similarity 19.4%; Pred. No. 9.2e-08;
 Matches 501; Conservative 303; Mismatches 887; Indels 890; Gaps 112;
 QY 4 GAARKOSTPG-----SLFLSPAPAPKAGSSSDS--VGEKLGAAALAAVYKRT 51
 DB 871 GAPKQAPAPBQPTSGQPKSTQAPAPAPAKSIPVKETAPAAEKLEPRAEQAPTVKR 930
 QY 52 EBYRRRTMTDKDSRGAATTTTTHRRFRSVCDSNATALEPLPLS-LPQPSIPAA 110
 DB 931 TETEKPPPI-KDSKSLTAEP-----QKAVL-----PTLKESPKRESTCPLCKTEIN 977
 QY 111 VPQAPPE-----PH-REETVATATG-----QVAGQPPAAABGE 145
 DB 978 IGSXDPHFNFNTCTECKQVCMCGNPTPHLTENCQTQRAISQGLGDIRKMPAPSGP-- 1035
 QY 146 QAVAGPASTVPSSTSKRPVSGPLVSGKEPPAPAGSGGSAKEQERSSQODDIE 205
 DB 1036 --KSPMPVPTSSSQKTAVPQVYL-----KKQDEVTAEKYL 1076

QY 206 ELKTAAGM-----SNDGRFLKFDIEIGSGSFXTVYKGLDTETTVAVMGELODRKLT 258
 DB 1077 EYKKEKTLSEKIPVAVTTDQKQESKLEKDKAS-----ALQKKKPL 1117
 QY 259 KSERGRFEEAMKLGLOHPNIVRFYDSWESVYKGCICVLVTELMSTGLTKYLRFKY 318
 DB 1118 PEKKLIPEEKIRSEKKP-----LLEKKPPEDK--LLEPKATSA-P-----EE 1162
 QY 319 MKIKVLRWCROILKGLQPLATRPPIIHRDLKCNIPITG-PTGS-----VKIGDIGLAT 373
 DB 1163 QKHDLKS-----QVQALAEELERVAHPKTYQEKQQTKEGHPSGHPQSLPKREDVXTKT 1219
 QY 374 IRRASPAKSVIGTPPAPAPMEYEEKYDESVDYATACMCLMNAATSEYIPSECONAAQI-- 431
 DB 1220 IKEQP-----QPPCTAKPD--QEKEDDKSDTSSSQQPKSPQGLSTPTGSSDGISSLBE 1271
 QY 432 -----YRRTSGYKPAFPDKVAIPE-----VKETIEGCIKONKDERYSIKD 472
 DB 1272 IFSLIPTDEKDLIKGKQDSFSQESSPSSDLAKLESTVLSLEQAAGTLADEK-SEKK 1330
 QY 473 LNHAFQREGTGVRYELAEEDGKEI--AIKMLRIEDIKLKGKYKDNALIEPSPD 527
 DB 1331 TQPH-----EVSPEQPKQEKQSLSETLEITISEEIKESQERRDT--FKKD 1377
 QY 528 LERDVEDVAQEMWESGYVCEGDHKTMAAIDRVSLIKRRQRQVNEBOEK-----582
 DB 1378 SQQDIPS-----SKDHKEKSEFVD--ITRRBPYDVSSESSSESSPVP 1420
 QY 583 -----KQESS-----LKQVEQS 596
 DB 1421 QRRRTSVSSSSDEYKQEDSQSGEEDFIRQIIEMSADEASGEDEFIRNQLKEI 1480
 QY 597 SASQTIKQLPSASTGIPTAS-----TTASAVSTOVE-----DEEPA-DHQOQLOV 642
 DB 1481 SSTBQKKEFTGKKGKXITAGKRRLTRKSSSTIDDAGRHSHWDEDAFPESPRTKY 1540
 QY 643 QQPSISVLSDPTVDSQGGSVF-----TSSRV 669
 DB 1541 RETKSQSEELVYTGGGGLRFTIETINSTIADKYSASSQKTSLYFDEPELEMSLT 1600
 QY 670 SSGQTYSGQHQAISTGVNPHIPSTV-----QAOSQPHV 707
 DB 1601 DSEBDSRGSGSSSLHSAFTPTGTSVSLDSDSSSHKKGESQKQKARHRHP 1660
 QY 708 YPSSVAQO-----SQQPSSS-----LT 728
 DB 1661 LPLTIBDSSEELREHEELLKQEKORIEQQOKSSKSKXKDELARRRERPKT 1720
 QY 729 GVSSSQPIQHPQOQCGIQQTAPRQQTQVYLSQSTSTSEATTAQVPVQAPQVLPQVSA 788
 DB 1721 PPSNLSPIEDASPTPEELRQAAMEELRRSSCSEYSPSE--SDPEGEFISPEKILTEVOK 1777
 QY 789 GKQSTQGVQVAPAEVAVANQPOATPTLASSVDSASHVASGMSDGNENVPSSGRHE 848
 DB 1778 VYKLPVAVSLYSPTEDSIQKESQ-----KAKSAEEMYE 1814
 QY 849 GRITKRYKRSVRSRHEKTSREPKRLILVSNKGRVVECO-----LETHNRKRV-- 899
 DB 1815 EMHMKH-----KYAPFAANRDEVEFEKEPLYGMLLEDYIESLVE 1857
 QY 900 -TFKFDIDGD--NPEIATITMNDLTALIERSPVDQVREILEKXDEMISEVSEPE 955
 DB 1858 DTNGSVAGSLTRQOE-----ENGPMQOKGRQCK-----RLSEQIYEDP- 1898
 QY 956 GDQGLESLQGDVDFGFSGQKLEGEFPQOPFASMPQOIGITSLITQVHSAQRPIVS 1015
 DB 1899 -----MOKITD-----LQKEFYELBSLHVQEDLVSS-----FL-- 1930
 QY 1016 VPESRLRESKYVPSKITD--TVAASTAQSPGMULSHSASSLSIQQAFSEL--RRAQMT 1070
 DB 1931 -IPESH-----ELVDLGTWVSTEEBKLT-----LDADAAYEELMKQOMQOLT 1972
 QY 1071 EG--PNTAPNFSHTGPT--FPVP-----PLISSTAGVTTAAATAPVATSPSPDI 1120

```

Db 1973 PGSSTQAPIGEDMESTMDPRMDASLTSSVLSGASLTSTSSATLSIPVKTQH-F 2031
QY 1121 STSVLOS-E-VTPTEGIAGVATSTGVVT-----SGGLPIPVSSPLSSVSSI----- 1170
Db 2032 STBEIEDYVDYTHREIGETIAHESLITLTPSSSAISVPP-SPTPLSTSSVSVCTTD 2090
QY 1171 -----TTPAVSISTSPSLQVPTSTSEIVVSTAL-----YP-----SVVSAISASAG 1215
Db 2091 SSSPITLITLTYTVEPDMITKPEDESEISSYFPFGSIIDYPEELSVLDRAPBDG 2150
QY 1216 GSTA-----TPGPPRAVVS-----COAGSTVGAULT 1244
Db 2151 RASADHVIYISLDMASSTIESVPPREGPVADVTDLISEKDPVKAKKETGNGITILE 2210
QY 1245 SVSTTSPSTASQSLTSSSTSTPTLAETV-----VSAHSLKTHSSST-----GLAF 1236
Db 2211 VLE-----AYDKKELEAKRKSLSLSEIVFHPSSVIALPMKEQLSTYFTSGETF 2262
QY 1297 SLAPSSSSSSPGA-GVSSYISQPGGLHPLVPSVIASTPLPQAAGPTSTPLPQVPSIP 1355
Db 2263 GOEKPAQOLPSGSPSVSISLPAPKPPFPSSSLDISAQP-PPPPPPPPPPPPPPPPPP 2320
QY 1356 PLVGPVANNVAVQOTLIHSOPALLPNQPTHCPEDVSDTQPAAPGIDIKYLEKLS 1415
Db 2321 PLPPP-----TSEKPTILPKKLT-----VASPVTAIPLFDVAT----- 2356
QY 1416 LFSEHSSSGAASHASVLSSTLVISVTPGIPPTAVAPSKLSTSTSCIPNTLPIGTV 1475
Db 2357 -----LETTAVLRSN--GLPVTRIC-----TTPAP----- 2379
QY 1476 ALPPTPVATGQVSTPVSTTSGV-----KQGTAPSKPPLTKAPVLPVGTLPAGTLPSQ 1531
Db 2380 --PVP-----KPSISPSGLVFTHRP--EPKSPPIAPKPIYI--QLPTTQKPTD 2423
QY 1532 LPPEP-GPSLTQSGQPLEDDAQRRLTSEKXITVTA-----VGPVMAAPTAITEAGTQ 1586
Db 2424 IHKPTGLSLTSM-----TLN-----LVTSADYKLPSPSPSLPHS-----NKSS 2464
QY 1587 POKGVQVKEGPVLANSSGAGVFMGRFQVSVADAGQKEGKNKSEAKSVHSESTSES 1646
Db 2465 FRFSKSLTEHYVITLPSERG-----PTDSAS 2493
QY 1647 SVUSS--SEPESTLYKPEPNGITIGISS-DVPESAKHTTASAKSDTQOPTV--GR 1699
Db 2494 QALTWPLGSPKDLVSEPFVPEPVTAIEIPISGEQFYISGALQTSAPVTAAPS 2553
QY 1700 FCVTTANKVGRF--SVSTEKIDTDEKEG-----VASPPMLEQAVLPAVIT 1748
Db 2554 FOAAPT--VTQPLTTEVSKTE--VSATKSTASVGLSISITIPP--EPLALDNI-- 2603
QY 1749 KKEKPELSEPHLNGSPSEAFALSRDVGSGSPHSPHOLSKSLPSONQSLSNSF 1808
Db 2604 HKEKPYKEDGKQI-----LVGDVILRTVKEVKTIDKICIDLSASTMDYKQI 2653
QY 1809 NSSYMSDNESIEDDLKLELRRLDKHKIKIQLDSQKHEIESLYTKLGKVPVAVII 1868
Db 2654 TANEVYGKOISAVOPSIINLSV-----TSLIVPVSLATEVTIF 2692
QY 1869 PPAAPLSGRRRPTKSGKSSSSSLGKSPOLSGNSGQSAASVLAHQOTLHPGNIP 1928
Db 2693 VVCT-----ASHSYTGTES--LVGAHEAMTTPLOLTSTKAEBEYRIP 2734
QY 1929 ESGQNLQPLRSPSSDNLYSAFTSDGA-----ISVPSIAPGQSGSTNTVAT-- 1979
Db 2735 --SDQVPIAREEARINLSIGTPAHAVTALITKVTYPPPGVYNGWTDSTVSQCIDTG 2790
QY 1980 -----VNSQAQAQPPA-MTSSRKGTETDLAKLV----- 2008
Db 2791 EYVDLSTKSHRTVYTMDESTSSVMTKILEDEKPVDLTGRBAVCCDVYKLPFGSCTA 2850
QY 2009 -----DNWARMNLSGRG-----SKGHMNYEGGGMARK 2038

```

```

Db 2851 QOPATLPEDEFGYRDDHYQYDRSPGYRGIGCKKPSMSDTNLAEGHFFYKSN-AFD 2909
QY 2039 FSAPQOLCISMT-----SNUGSAP-----ISAASATSLGHFTSMCP--Q 2078
Db 2910 YSEGTDFVADLTSGRVTGTEVWDYSSKRTGYPPEIRQYISGAGISTPOYSTARKMPPPG 2969
QY 2079 QY---GFPATPFQAGMSTGAPAPQPLGQFQPVGTASLQNTNISMLQKISNPPGSNIRT 2135
Db 2970 QYCVGSVLRSSNGVSVVATPTPTSTF-----AITTQPSIFST 3008
QY 2136 T 2136
Db 3009 T 3009

RESULT 7
MAPB_HUMAN
ID MAPB_HUMAN STANDARD; PRT; 2468 AA.
AC P46821;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 1B (MAP 1B) [Contains: MAP1 light chain
DE LCI].
GN MAP1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain.
RX MEDLINE=95104835; PubMed=7806212;
RA Lien H.L., Feener C., Fischbach N., Kunkel L.M.;
RT "Cloning of human microtubule-associated protein 1B and the
RT identification of a related gene on chromosome 15.";
RL Genomics 22:273-280 (1994).
CC -!- FUNCTION: The function of brain MAPS is essentially unknown.
CC phosphorylated MAP1B may play a role in the cytoskeletal changes
CC that accompany neurite extension. Possibly MAP1B binds to at least
CC two tubulin subunits in the polymer, and this bridging of subunits
CC might be involved in nucleating microtubule polymerization and in
CC stabilizing microtubules.
CC -!- SUBUNIT: 3 different light chains, LCI, LC2 and LC3, can associate
CC with MAP1A and MAP1B proteins.
CC -!- DOMAIN: Has a highly basic region with many copies of the sequence
CC KKKK and KKEI/V, repeated but not at fixed intervals, which is
CC responsible for the binding of MAP1B to microtubules.
CC -!- PTM: LCI is coexpressed with MAP1B. It is a polypeptide generated
CC from MAP1B by proteolytic processing. It is free to associate with
CC both MAP1A and MAP1B. It interacts with the amino-terminal region
CC of MAP1B (by similarity).
CC -!- SIMILARITY: TO MAP1A.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; L06237; AAA18904.1; -
Genew; HGNC:6636; MAP1B.
DR MIM; 151129; -
DR GO; GO:0005875; C:microtubule associated complex; TAS.
DR InterPro; IPR000102; MAP1B_neuraxin.
DR Pfam; PF00414; MAP1B_neuraxin; 10.
DR PROSITE; PS00230; MAP1B_NEURAXIN; 6.
KW Microtubule; Repeat; Phosphorylation.
FT CHAIN ? 2468 MAP1 light chain LCI.
FT REPEAT 1878 1894 MAP1B 1.
FT REPEAT 1895 1911 MAP1B 2.

```

FT REPEAT 1912 1928 MAP1B 3.
 FT REPEAT 1929 1945 MAP1B 4.
 FT REPEAT 1946 1962 MAP1B 5.
 FT REPEAT 1963 1979 MAP1B 6.
 FT REPEAT 1997 2013 MAP1B 7.
 FT REPEAT 2014 2030 MAP1B 8.
 FT REPEAT 2031 2047 MAP1B 9.
 FT REPEAT 2048 2064 MAP1B 10.
 FT DOMAIN 589 790
 LYS-RICH (HIGHLY BASIC, CONTAINS MANY KEE AND KKEI/V REPEATS).
 SC SEQUENCE 2468 AA; 270618 MW; 540839CDBF09D461 CRC64;
 Query Match 4.1%; Score 439; DB 1; Length 2468;
 Best Local Similarity 20.6%; Pred. No. 8.4e-08;
 Matches 452; Conservative 271; Mismatches 828; Indels 640; Gaps 102;
 QY 90 ATALEPLGLPLSLPPSI-----PAA--VPOGAPPEPHEETVATATAT 131
 DB 516 ATQKDLTG---QVPTFVVKQTKLKORADRESILKPAKPLPSKVKESKEETPEVTATYN 572
 QY 132 QVAQCPAAAARCEQAVAGPAP---STVSSSTKDPVSGPSLVGSEPPPARSGSG-- 186
 DB 573 HV-EKPPKVESEKWKVKKDKPVKTETKTSVTEKEVP-----SKEEPSVKAFAVEK 623
 QY 187 -----GGSAGE---POERSQOQDDIEELE--TKAVGMSNGRFLKFD----- 224
 DB 624 QATDVFKPAKKEKTVKKTETKVPEDKKEKPKKXVAKKEDKTIPIKKEKPKKEVKE 683
 QY 225 --LEIGSGFKTVYKGLDTEETVEVAMCEL--ODRKLTSEKORFKEEAMKGLQHPNI 280
 DB 684 VKKEIKKEEKEPKKEVKEKTPPEKVEKVEKKEKKEKKEKKEKKEKKEKKEKKEKKE 735
 QY 281 VREYDSMESIVKGGKCLVLTETMTSGTLKTYL--KRFKMKIKLRSWROQLKGLQFL 338
 DB 736 -----PKDAKSSSTPLSEAKKPAALPKVPKKEESYKKSVAAGKPE--KG--- 780
 QY 339 HTRPPIIHDLCDNFI---TGPTGSYKIGDGLATL--KRAFSKAVIGTEPFMA 391
 DB 781 ---KIKYIKKEGKAAEVAAVAGTATTAVMAAGIAAGPAKELAEBSLMSPPDLT 837
 QY 392 PEMYEKYKDSVDYAFGCMLEMATSEYPSSECQ---NAAQIYRVTSGVKPPASFDKY 447
 DB 838 KD-FEELKAEVDTKQIKPQLLEIIDEKLIKETEPEAVYIQRERVTKG--PAESPDE 894
 QY 448 AI-----PEKKEIIEGCIKONKBERYSIKLNLHAPFOE--ETGVAVELAE-- 491
 DB 895 GILTTEGEGCEQTPBELEPYE---KQGVDDIEKFD--EGAGFEESSEEGDYEEKATE 949
 QY 492 -----EDDGEKIAIKMLRIEDIKLKGKYNKONPAIEFSPLEDERVEDVAQEWESGYV 546
 DB 950 EAEPPEEDGE-----EHVCVSASKISPTED-EESAKAEADA---YIREKRES--V 993
 QY 547 CEGDHT---MAKAIKRVSLIKREROROLVREEOK---KQESSILKQVQ--- 595
 DB 994 ASGDPRABEDMDAIE-----KGEAQCSEBEADEDEKAEDEAEFEYEPMEAEEDYMA 1047
 QY 596 --SSASQGTG-----IKOLPSASTG-----IPTASTTASVSTQVPE 630
 DB 1048 VVDAARAEGAEBOYGLTTPTKQIGASPGREPASSIHDTLLGSESEKTAJ---DE 1103
 QY 631 EPADQHQOQO---YOOPSIISVSDGT-----VDSGQSSSVFTESRVSQQTVS 676
 DB 1104 ENREDQBEFTATSGTYSTIEISSEPTPMDEMSTPRDWMSDENNETES--PSQEVFN 1161
 QY 677 YGSGHEQAHSTGTGPHIPSTVQAQSQPHGYVPPSSVAQSQ---SQGQSSSSLTGVSS 732
 DB 1162 ITKXESSLYS-----QEYSKPADVTPLNGFSESGSKTATDCKDNANASTISIP 1209
 QY 733 SQPIQHPOQOQIGQGTAPQGT-----VOYSLQSTSTSEATTAQVSGOPAPQVLPQ 785
 DB 1210 PSSMEDEKFSRSALRDAYCSEKVASSTLIDKDISIAVSEKVSPEKSPSLSPSPSPLEK 1269
 QY 786 VSAGKOS-----TQGVSGVAPABAPVAVAQPATQPTT---LASVDSASHSDVAGSMGSDGN 837

DB 1270 TPLGRSVNFSILTPNEIKVSAEAFAVPSPEVTOVEEHCHSPDKTLELVSP-----S 1324
 QY 838 ENVPSSGRHREGRTTKRHYKRSVRSRSRHEKTSREKRLILNVNSKNGDRVVECOLETHNKK 897
 DB 1325 QSVTSAGH-----TPYQSTPDKSKSHLPTE-----VIEKPAV 1359
 QY 898 WTFKFDLDGDNPEBIATIMVNNDFILAIERSEFVDQVREIIEKADEMISEVSVPEEGD 957
 DB 1360 PVSFEFS-DACNENRASVSPWDEPV--PDSESPLEKVLSPRSPPLIGSE----- 1407
 QY 958 QGLESLQGDYDYGEGSQKLEGEFQPLPASMPOQIGTPSSLTQV---HSGRRFI 1013
 DB 1408 SAYESFLSADDK--ASGRGAESPFEKSGKQSGPPQVPS-VSEMTSTSLYQKQEGKSTD 1464
 QY 1014 VSPVESRRLRESKVPSEETDVAASSTAQSPGMNLSH-----SASSLSLQO--AFSELR 1065
 DB 1465 PAPIKEDFGQEKK-----TDVEMASSQ-PALALDERKLGQVSPQIDVSGFGSKEDT 1517
 QY 1066 RAQMTGEPNTAPPNESHTGPTPEPVVPPFLSSIAGVPTTAATAPVPAISSPPNDISTSVI 1125
 DB 1518 KMSISEG-----TV 1526
 QY 1126 QSEVTVTEEGIA-----GYAT-STGVNYSGLPIPPVE-SPVLSVSVSITIPAV 1175
 DB 1527 SDKSATPVDEGVAEDTYSHMEGVA SVTASVATSSFPETTTDVSPLHAEGS----- 1580
 QY 1176 VSISTSPLOYPTSTSEIYVSVSTALYPVATVATASAGSSTARGPPRPVAVSQQAAG 1235
 DB 1581 -----PHSTVEDLSISVAVQPTTFQETEMSPSKBEC-----PRP----- 1616
 QY 1236 STTVGATLTVSSTTSFPSTASQSLTQSSSTPTPLAETVAVSAHSLDKTSHSSTTGLA 1295
 DB 1617 -----MSISPPDPSPKTA-----KSRTPV-----QHRSEQSSMSIE 1648
 QY 1296 FSLSAPSSSSSPAGVSVSISQPGGLHPLVTSVIASTPILPQAAGPTSPILLPO----- 1350
 DB 1649 FGQESPEQS-----LAMDFSRQSPDHPPTVAGACVLIIT-----ENGTEVDYPSDMQDS 1697
 QY 1351 --VPSIPLVOPVAVNPVAVQOTLIHSOPALIPNQCPEVSDTOPKAPGIDDIKT 1408
 DB 1698 SLSHKIPMEBEPYTOQNDLISELI-SVSQVEASPSSTSNATP-----SQASLQGED--T 1749
 QY 1409 IBEKRLSPFSEHSSGQAQASVSLSTGLVSTVTPGIPTTAVAPSKLTLSTSTCLPPT 1468
 DB 1750 L-----SDVAPPRDMSIYASLSE-----KVOSLEGEKLSPKS 1782
 QY 1469 NLPLGLVALPVTVVTPPGOVSTVSTTSVTKGCTA---PSKPLTKAPVLVGTETL-- 1522
 DB 1783 DIS-----PLTPRESPLVPTFSSTSAVKERTATCHSSSSPPIIDMAABEYGFRAV 1836
 QY 1523 -----PAGTLPEBOLPPFPGPSLTLTOSQOPLLEDIDAQLRRTLS 1559
 DB 1837 LPTMQHHLALNNDLSTPGLEKXSGKTPDPFSYAYAKPEETTRSPBEEDYDE-----S 1891
 QY 1560 PEKITYTSAVGPVSMAPTAITAGTOPQGVQVKEGYPVLATSSGAGVTRKMGHFOVVA 1619
 DB 1892 YEKTRTSVDG-----GYEYK-IERTKSP--SDSGSYETIHK-TTKTP 1933
 QY 1620 ADG-----AQEGKNSDEAKSVHFPSSSTSSSVSSSPEST-----LYKPEPNC-- 1665
 DB 1934 EDGYSYEIIEKTRTPEEGYSYDISEKTSPEPVSGYETEKERSRRLIDISNGYD 1993
 QY 1666 -----ITIPGISDVPESAHKTTASEANS-DTQPTTKVGRFOVTT-----TANKYGRF 1712
 DB 1994 SEDGHTLGPDSYVETTEKTSFPEBEGSVETS--TKTRTQDTSTYCYETALEKILRT 2051
 QY 1713 SVS-----KTEDKITDTKKGVPVAPPPMDLEQAVLPAVLPKPKKPELSESHLNGSSD 1767
 DB 2052 PQASTYVETSTDLCTYTEKSP--SEARQVDLCLVSSCEYKHKPRLS--PSFIN---PN 2105
 QY 1768 PEAFLSRDUDVDSGSP-----HSPHQSSSLSPSONLSQSLSN 1806

DB 514 SYPGIGKV-----DIIATQONDGNGQGPVAGVNSST 546
 QY 852 TRHRKXSVRSRSHKXTSRPLRLINNSKGDVVECOLETHNRKMTFKFDLDGDNPE 911
 DB 547 S-----STASAK-----DRP- 556
 QY 912 EIATIMVNNDFILATERESFVDVVEIIEKADEMUSEDEVSEPEGGDGLSLQKDDYGF 971
 DB 557 -----LSARR-----RLTKSGEEMLPSPGAVQ-RTPSAVEPLKPOE----- 594
 QY 972 SSGQKLEGEFKQPIPPASMPQOIGIPSTLTQVHSAGR--RPIVSPVPSRL----- 1022
 DB 595 -----DDPIQAFRSSD-----CSITQNMHTLPREKEXRLMHGLSEDELSSSTSS 639
 QY 1023 -----RESKFPSEITDYTAAGTAPGKMLSHSASSLSLQCAFSLR 1065
 DB 640 TDKSDGSRGKSHTEKMDLVQMLT-QTLRLKAKESCEDLOVLNPGSEFR 689

RESULT 9
 MINT MOUSE STANDARD; PRT; 3644 AA.
 AC 062504; Q80TN9; Q99PS4; Q9QW2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE MSX2-interacting protein (SMART/HDAC1 associated repressor protein).
 GN MINT OR SHARP OR KIAA0929.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE OF 49-3644 FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR
 RP LOCATION, TISSUE SPECIFICITY, DNA-BINDING, AND INTERACTION WITH MSX2.
 RC TISSUE=Testis;
 RX MEDLINE=93379811; PubMed=10451362;
 RA Newberry E.P., Latifi T., Towler D.A.;
 RA "The RRM domain of MINT, a novel msx2 binding protein, recognizes and
 RA regulates the rat osteocalcin promoter.";
 RA Biochemistry 38:10678-10690(1999).
 RN [2]
 RP SEQUENCE OF 1-112 FROM N.A.
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
 RA Hachikume W., Hayashida K., Hirozane T., Hori F., Imotani K., Kono H.,
 RA Ishii Y., Itoh M., Kagawa I., Kawai J., Kojima Y., Kondo S., Kono H.,
 RA Koya S., Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,
 RA Ohno M., Ohsato N., Saito R., Sakazume N., Sano H., Sasaki D.,
 RA Sato K., Shibata K., Shiraki T., Tagami M., Takeda Y., Waki K.,
 RA Watanuki A., Muramatsu M., Hayashizaki Y.;
 RA Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 69-3644 FROM N.A. (ISOFORM 2), AND VARIANTS THR-348;
 RP PHE-762; PHE-773 AND LEU-933.
 RC STRAIN=ICR; TISSUE=Brain;
 RA Satamoto T., Gotou T., Isagawa Y., Mimura H., Kimura K., Kawaihi M.;
 RA "MINT/spen negatively regulates Notch signaling by inhibiting RBP-
 RA J/5b(H) activity.";
 RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 318-578 FROM N.A.
 RC TISSUE=Cochlea;
 RX MEDLINE=97237053; PubMed=9119401;
 RA Ciozet F., El-Amraoui A., Blanchard S., Lenoir M., Ripoll C., Vago P.,
 RA Hanel C., Fiazmes C., Levi-Acobas F., Depietis D., Mattei M.-G.,
 RA Weill D., Pujol R., Petit C.;
 RA "Cloning of the genes encoding two murine and human cochlear
 RA unconventional type I myosins.";
 RA Genomics 40:332-341(1997).
 RN [5]
 RP SEQUENCE OF 2598-3644 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;

RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT 11. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RT DNA Res. 10:35-46(2003).
 RN [6]
 RP TISSUE SPECIFICITY:
 RX MEDLINE=22261914; PubMed=12374742;
 RA Oswald F., Kostetzka U., Astrahantseff K., Bourteele S., Dillinger K.,
 RA Seidner U., Ludwig L., Wilda M., Hameister H., Knoechel W., Lipay S.,
 RA Schindl R.M.;
 RT "SHARP is a novel component of the Notch/RBP-Jxappa signalling
 RT pathway.";
 RT EMBO J. 21:5417-5426(2002).
 RN [7]
 RP FUNCTION, AND TISSUE SPECIFICITY:
 RX MEDLINE=22483652; PubMed=12594956;
 RA Kuroda K., Han H., Tani S., Tanigaki K., Tun T., Furukawa T.,
 RA Taniguchi Y., Kurooka H., Hamada Y., Toyokuni S., Honjo T.;
 RT "Regulation of marginal zone B cell development by MINT, a suppressor
 RT of Notch/RBP-J signalling pathway.";
 RT Immunity 18:301-312(2003).
 CC -1- FUNCTION: Essential corepressor protein, which probably regulates
 CC different key pathways such as the Notch pathway. Negative
 CC regulator of the Notch pathway via its interaction with RBP5UH,
 CC which prevents the association between NOTCH1 and RBP5UH, and
 CC therefore suppresses the transactivation activity of Notch
 CC signaling. Blocks the differentiation of precursor B cells into
 CC marginal zone B cells. Probably represses transcription via the
 CC recruitment of large complexes containing histone deacetylase
 CC proteins. May bind both to DNA and RNA.
 CC -1- SUBUNIT: Interacts with NCOR2, HDAC1, HDAC2, RBBP4, MBD3 and
 CC MTAIL1. Interacts with the nuclear receptors RAR and PPARG.
 CC Interacts with RAR in absence of ligand. Bind to the steroid
 CC receptor RNA coactivator SRA (by similarity). Interacts with MSX2.
 CC Interaction between RBP5UH and NOTCH1.
 CC -1- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q62504-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q62504-2; Sequence=VSP 008564;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Highly expressed in testis. Expressed at lower
 CC level in brain, lung, spleen, liver and kidney. Weakly expressed
 CC in cardiac and skeletal muscles and ovary. In spleen, it is
 CC expressed in follicular B-cells, while it is weakly expressed in
 CC marginal zone B-cells.
 CC -1- DOMAIN: The RID domain mediates the interaction with nuclear
 CC receptors.
 CC -1- DOMAIN: The SPOC domain, which mediates the interaction with
 CC NCOR2, is essential for the repressive activity (By similarity).
 CC -1- SIMILARITY: Belongs to the Spen family.
 CC -1- SIMILARITY: Contains 1 RID (receptor interacting) domain.
 CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
 CC -1- SIMILARITY: Contains 1 SPOC domain.
 CC -1- CAUTION: Ref.4 sequence differs from that shown due to multiple
 CC framehifts and conflicts that create stop codons.
 CC -1- CAUTION: Ref.5 sequence differs from that shown due to what seems
 CC to be the presence of intronic sequence in the cDNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: BY26481; -; NOT_ANNOTATED_CDS.
 DR EMBL: AF156529; AAD55931.1; ALT_INIT.
 DR EMBL: AB055980; BAB32786.1; -.
 DR EMBL: Z78160; CAB01562.1; ALT_SEQ.
 DR EMBL: AK122402; BAC65684.2; ALT_SEQ.
 DR MGD: MGI:1891706; MInt.
 DR InterPro: IPR00504; RNA_rec_mot.
 DR Pfam: PF00076; xtm; 3.
 DR SMART: SM00360; RRM; 3.
 DR PROSITE: PS00102; RRM; 4.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 DR PROSITE: PS00917; SPOC; 1.
 KW Transcription regulation; Repressor; Nuclear protein; DNA-binding;
 KW RNA-binding; Repeat; Coiled coil; Alternative splicing; Polymorphism.
 FT DOMAIN 1 574
 FT 6 81
 FT 336 416 RNA-BINDING (RRM) 1.
 FT 439 514 RNA-BINDING (RRM) 2.
 FT 518 590 RNA-BINDING (RRM) 3.
 FT 559 575 RNA-BINDING (RRM) 4.
 FT 822 850 COILED COIL (POTENTIAL).
 FT 1185 1206 COILED COIL (POTENTIAL).
 FT 1509 1544 COILED COIL (POTENTIAL).
 FT 1607 1627 COILED COIL (POTENTIAL).
 FT 2216 2704 COILED COIL (POTENTIAL).
 FT 3478 3644 RID.
 FT 2138 2462 SPOC.
 FT 2706 2845 INTERACTION WITH MSX2.
 FT 125 277 INTERACTION WITH RBPSUH.
 FT 236 326 ARG-RICH.
 FT 648 721 SER-RICH.
 FT 702 832 TYR-RICH.
 FT 2101 2233 ARG-RICH.
 FT 2377 2518 ALA-RICH.
 FT 2950 3475 PRO-RICH.
 FT 618 640 PRO-RICH.
 FT 348 348 Missing (in isoform 2).
 FT 762 762 /FTId=VSP_008564.
 FT 773 773 I -> T.
 FT 933 933 S -> F.
 FT 754 754 S -> L.
 FT 1524 1524 R -> G (IN REF. 3).
 FT 1560 1560 D -> A (IN REF. 3).
 FT 1570 1570 H -> Y (IN REF. 3).
 FT 1574 1574 F -> L (IN REF. 3).
 FT 1609 1609 R -> G (IN REF. 3).
 FT 1659 1659 Q -> R (IN REF. 3).
 FT 1669 1669 I -> V (IN REF. 3).
 FT 1705 1705 S -> F (IN REF. 3).
 FT 1815 1815 V -> A (IN REF. 3).
 FT 2097 2097 A -> V (IN REF. 3).
 FT 2201 2202 G -> A (IN REF. 3).
 FT 2322 2322 MISSING (IN REF. 3).
 FT 2385 2385 A -> V (IN REF. 3).
 FT 2502 2502 P -> Q (IN REF. 3).
 FT 2505 2505 R -> Q (IN REF. 3).
 FT 2519 2519 E -> K (IN REF. 3).
 FT 2554 2554 D -> N (IN REF. 3).
 FT 2679 2688 T -> S (IN REF. 3).
 FT 3010 3010 L -> P (IN REF. 3 AND 5).
 SQ SEQUENCE 3644 AA; 398750 MM; 9C7EC49A81A7D4A4 CRC64;

Query Match 4.0%; Score 429.5; DB 1; Length 3644;
 Best Local Similarity 19.4%; Pred. No. 2,7e-07;
 Matches 443; Conservative 286; Mismatches 910; Indels 649; Gaps 98;

DB 62 DKDSGAATTTTTHRRFRRSVIT-CDSNATALEPGLPLSPPSIPAAVQASAPPEFH 120
 1037 DTSRDGVLIRGSESRKVRKELLRKESKYLRLNSALSPKDCQDPAAVAGSGSNRS 1096
 QY 121 RE-----ETVATNTSGVAQCPAAAPAGEDAVAGPAPSTVP--SSTSKRPVSPSLV 172

DB 1097 SDVHAGLGEIETHGVSVEIQTETO--PKKALPSK-----PPKQOLLENOGPEKEEVRKNYC 1149
 QY 173 GSKKEPPPARSGS---GGGSAKEPPEERSQOODDIEELTEKAVGMSNDGRFLKPIEIEGR 229
 DB 1150 RPREEPAPHRAGQEKPHGNAHE-----KGLDIDHQSIRKQMEGRRKQRMEMETAK 1203
 QY 230 G-SPTVYKGLD-----TETT-----VEVAMELQDRKLTKS 260
 DB 1204 AEKFGSPKKDDVDYERSLVEHVGKPPQDVIDDSDPPSKRRRTDHDVDFICTKREKNY-RS 1262
 QY 261 ERQRFKEAEML-----KELQHPNIV----- 281
 DB 1263 SRQ-ISEDSSEITSSPSVRHSGFHDHDDPGSPPLVSVKSGPKGDEKGLPYNAVRDDP 1321
 QY 282 ---RFDYSESTVYKGCIVLVTEIMTSGILTKYLKFKWIKIKILRS-----WCRQIL 332
 DB 1322 LKCPNYS-----GKR-----EGTADTK-----IKLVLNSEGSPRW----- 1355
 QY 333 KGLQFLHTRTPPIIHRDLKCDNITITGPTGSKTIGDGLATLKRAAPKSVIGTPEPMAP 392
 DB 1356 -----DPEMKQDSRDP--VSFPN-----SVTKHDSLRKRSVRDLE--P 1390
 QY 393 EMEYEKYDESVDVYAFGCMLEMATSEYPSYEQNMAQIV----- 432
 DB 1391 GEVPSDSDDEAR-----HRSQPRASFTDSPPLSLFLDRDQKLRERD 1434
 QY 433 RRYTSGYKPAASF-----DKVAIPVEKKEIIEGC--IRQNKDERYSIKDL-INHAFQCEETG 484
 DB 1435 ERLASSLERNKFFYSFALDKTITPDTKALIRAKSLSSSRRENNFLLWDCRFANFRNNKD 1494
 QY 485 VRVELAEEDDEGEKIALKLMIRIEDIKLK-----GKYKNEALIEFSLERDVEDVAQ 538
 DB 1495 -----KEKUDSAPRLPSWY---MKKKKLRITDSEGLDDDKD-----ERREEOERQ 1538
 QY 539 EMEVSGYVCEGDKHTAKAIKDRVSLIKRKEQ----- 571
 DB 1539 ELFASRFL-----HSLTFEODSKRLQHLERKSESDPPRGLYGRQMSGANSTSDSVQEP 1594
 QY 572 -----RQVAREDEQEK-----KQESSLSIKQVEQSGASQGTGKQLPASSTGI 613
 DB 1595 VVLEHRSFMEITMOQKEKEKQPKAEAKQKEPEETHPKPEPAALTEKEPEKAPVAGL 1654
 QY 614 PTAST-----SASVSTQVE-----PEEP-EADQHQQLQYQOPSISVLDGT 654
 DB 1655 PAVTIVTTPPEPSSAPKEAEAEAPASPAGEKAPAPVSEETKLVSEVPVAPQPRQ 1714
 QY 655 VDSGQSSVFTESRVSSQQTVSYSQHEQAHSGTGPBGHLPSTVYQAOQSGHGYPPSSVA 714
 DB 1715 SVPPPE-----DSRDSQDSALALAPAPQESAATDAVF-----CVNAEP---LTPGTYS 1761
 QY 715 QGSGQGPSSSLTGVSSQPI-----QHPQOQGIQCTAPPQQTIVQVLSQTSSTSEAT 769
 DB 1762 QVESSVDPEP-----SSPQLSKLTQREBAEAG--KYEKDPTTSTPEPATQNGVAGS 1813
 QY 770 TAQPVQAPQVLPQVSA-----GKSTQGVSAVAAPVAVAQPAQTOPTTIA 819
 DB 1814 EAQPAASEV-EANPVAAKDKRTKNSKRSKTSVQAAAASVVEKVPTRKSEIRIDREKLR 1872
 QY 820 SSVDAHSDVASGMSDGNENVPSSSGRHEGRTTK-----RHYRKSVR--RSRH 866
 DB 1873 SSSPRGEAQKLLELMKEKIKRTASKSGSGDTEHPEBPLPSRSRRNRVRSVYATMTIH 1932
 QY 867 EKTSPKRLILNVSKGDRVVECCQLETHNRKNVTEKFPDLIDGNPEIATIMVNDPILAI 926
 DB 1933 ESRSPAKSPV-----EOPRVTEKRLERLEIQEAVV-----PP-----TT 1965
 QY 927 REESFVDQVRELIIEKADMLSEDSVVEPEQGLSLOGKDYDYGSGGQKLBGERKQIP 986
 DB 1966 PRRGPRPTTKRAEEDGHEHREKPEETPRAPAGWSPSSQSAAGGQGRKGRNEQKVE 2025
 QY 987 ASSMQQIGIPITSSILTVVHAGRRFIVSPVESLRBS-----KVPSSEI----- 1032
 DB 2026 AAA---EAGAGASITREGNPKSGREMAASEPRKDRDRDSTKSGPDTTFVVEVLEKRPPEK 2082


```

QY 1033 TDVAASSTQSPGMNLSHSASSLSLQAFSELRRAQMTG-GENTAAPNFSHTGPTFPVVP 1091
Db 2083 TYKSKRGKRRKSTRSGMDRAHQBSLEMAARAQQAADKRAAASPOSESPOKSGSS 2142
QY 1092 PRLSSAGVPTTAATAAPVATSPSPNDISTS-----VIOSEVTPTEBGA---GVA 1141
Db 2143 POLANNPADPDBAEESASASTAPEGTQLARQIETLEQAVONIAQLPEPSAAASKGTA 2202
QY 1142 TSTGVTS-----GLPIPPSEPVLSVSSVST-----IPAVSISTSP 1183
Db 2203 TATATASEEPAREHGHKPAHQASETELAAGISIIISDASGEPEFNSAGAPSPVPGQTPH 2262
QY 1184 S-----LQVPTSTSEIVSSSTALVP---SVTSATSAG-----GSTA 1219
Db 2263 REGMEFGLHEASGILETGTATESAPQVSAIDPPGSGADTKTRNGSDSVQEAAGSVA 2322
QY 1220 TPGPKPPAVVSOQAAGSTTVGATLTSVSTTSPPTASQSLSLSSSTSTPTLAETVVS 1279
Db 2323 EWTTPPKDKGRKOTTRRRKRNANKVVALTETRASEAEQ-----TQSEBPAAEATPAAT 2376
QY 1280 AHS-----LDKTSHTSTTGAFSLASFSSSS-----SPGAGVSYSIQP- 1318
Db 2377 PEAPQEEKSEKPPSPAECTFPDSKTPPAESLSQENSAAEKTCKAPVLPALPPLSOPA 2436
QY 1319 ----GGHPLVIESVIASTPILPQAAGPTSTPLLP--QVPSIPPLVQPVANVPAVOQTILI 1372
Db 2437 LMDDGQARKVKSIIIESDPVTPPSDSGIPPTIPLVTIAKLPPVLP-GGV- 2488
QY 1373 HSOPAPAL-----PNOPTHCEVDSDTQPKAGIDITLKEKLSLSEH--S 1421
Db 2489 HQSPRPVETWITROEPPRAQSTPSPALPDTK-----ASDMOTSSSTLAKIMDPKXVS 2543
QY 1422 SSGAQAASVLETSVLESTVTPGIPPTANAP-----SKULTSTTSTLP-----PMTLPL 1472
Db 2544 ATGVTSTSV--TTAIAEPVASAPLOQAPAPPCDPKHPLEGVSAANVADTQASBPV 2600
QY 1473 GTVALPTVPTVPTQVST-----PVSTTSQVQKQPTASKPLTAPVLPVTEP- 1522
Db 2601 AADKEKAPVIAPIAKITISVIRMPVSIULENSQKTTAKAPQITLGLVSLTGLVANSV 2660
QY 1523 PAGTLPS-----EQLPPFPSPSLTQSQOPLDLDAOLRTRLSEXT--TVTSAGV 1570
Db 2661 PVNALKGPVKSVALTKGLVSTPAGVNLKGPVNVLTGPVNVLTTPVSAITGVHNAAPG 2720
QY 1571 PVSMA-----AFTATTEAGTQ-OKG-----VSQVKEGPVLAISGA 1606
Db 2721 PVTAAACGVTAATGTAATVGAATAPAAKGKORASSNENSRRHPGSMVVIDRPA-DTSGGA 2779
QY 1607 GV-FKMGRFQVSVADGAKQKQK-----SEDAKSVHFESESTSESVLSSSSPES 1656
Db 2780 GLRVNTSEGVLLSYSGQKTGPPRIAKISQITPAPASMDIEFOQSVSKSVKXADS---I 2836
QY 1657 TLVPEPNGITIPGISSDVPESAHTTASAKSDTQPKVGFQVTTANKVGRFSVK 1716
Db 2837 TPTQSAKPGPTPBAFANV--AAHSTVLTAQTYNASP-----VASSYKTRDPSELEK 2886
QY 1717 TE-----DKITDTKKEGPVA-----SPPMDLQAVL--PAVITKKEKPELSESHLNGP 1764
Db 2887 PEPHILSVSTPVTQGTAVKVLTOGINTPVPVWNLVLTPLSIVTNNK--LADPVLTKIE 2944
QY 1765 SSDEBAEAFSLRDVDGSGSPHQSLSKSLPSQNLQSLNSFNFSVMSDNESDIDE 1834
Db 2945 TKVLQAPALGPTL-----TPHHPALPEK-LPA----- 2971
QY 1825 DLKLEIRLRDKHLKEIQLOSRQGHIESLYTKLGKVPVAVLP-PAAPISGRRRRPTK 1883
Db 2972 -----EVNHVPSGSETPADRTIAHLATPKDTHSRPRTGPTPGJPPRCH 3016
QY 1884 SKGSKSSSSSLGKNSPOLSGNLSCQASAAVLPQOT-LHPPGNIPES-----GQNO 1934
Db 3017 PSSITSTALST--NATVMAAGIWPQPFISISHPQSVIMPHSITQTVSLGHLGQGBVR 3074

```

```

QY 1935 LIQPLKPS 1942
Db 3075 MSTPLPS 3082

RESULT 10
PCLO_RAT
AC 09UKS6; Q9ULI1; STANDARD; PRT; 5085 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Piccolo protein (Multidomain presynaptic cytomatrix protein).
GN
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH RABAC1.
RX MEDLINE=20170257; PubMed=10707984;
RA Fenster S.D., Chung W.J., Zhai R., Cases-Langhoff C., Voss B.,
RT Garner A.M., Kaempf U., Kindler S., Gundelfinger E.D., Garner C.C.;
RT "Piccolo, a presynaptic zinc finger protein structurally related to
RT bassoon."
RL Neuron 25:203-214(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Fenster S.D., Cases-Langhoff C., Gundelfinger E.D., Garner C.C.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CALICUM-BINDING ACTIVITY, AND MUTAGENESIS OF ASP-4668, ASP-4674;
RP VAL-4688; MET-4689; VAL-4690; SER-4691; GLN-4692; ASN-4693 AND
RP ALA-4694.
RX MEDLINE=21181819; PubMed=11285225;
RA Gerber S.H., Garcia J., Rizo J., Suedhof T.C.;
RT "An unusual C(2)-domain in the active-zone protein piccolo:
RT implications for Ca(2+)-regulation of neurotransmitter release."
RL EMBO J. 20:1605-1619(2001).
CC -!- FUNCTION: May act as a scaffolding protein involved in the
CC organization of synaptic active zones and in synaptic vesicle
CC trafficking (By similarity).
CC -!- SUBUNIT: Interacts with Rabac1/Prai, RIMS2 and profilin (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Concentrated at presynaptic side of synaptic
CC junctions.
CC -!- ALTERNATIVE PRODUCTS:
CC Name=1;
CC Name=2;
CC Name=3;
CC Name=4;
CC Name=5;
CC Name=6;
CC Name=7;
CC Name=8;
CC Name=9;
CC Name=10;
CC Name=11;
CC Name=12;
CC Name=13;
CC Name=14;
CC Name=15;
CC Name=16;
CC Name=17;
CC Name=18;
CC Name=19;
CC Name=20;
CC Name=21;
CC Name=22;
CC Name=23;
CC Name=24;
CC Name=25;
CC Name=26;
CC Name=27;
CC Name=28;
CC Name=29;
CC Name=30;
CC Name=31;
CC Name=32;
CC Name=33;
CC Name=34;
CC Name=35;
CC Name=36;
CC Name=37;
CC Name=38;
CC Name=39;
CC Name=40;
CC Name=41;
CC Name=42;
CC Name=43;
CC Name=44;
CC Name=45;
CC Name=46;
CC Name=47;
CC Name=48;
CC Name=49;
CC Name=50;
CC Name=51;
CC Name=52;
CC Name=53;
CC Name=54;
CC Name=55;
CC Name=56;
CC Name=57;
CC Name=58;
CC Name=59;
CC Name=60;
CC Name=61;
CC Name=62;
CC Name=63;
CC Name=64;
CC Name=65;
CC Name=66;
CC Name=67;
CC Name=68;
CC Name=69;
CC Name=70;
CC Name=71;
CC Name=72;
CC Name=73;
CC Name=74;
CC Name=75;
CC Name=76;
CC Name=77;
CC Name=78;
CC Name=79;
CC Name=80;
CC Name=81;
CC Name=82;
CC Name=83;
CC Name=84;
CC Name=85;
CC Name=86;
CC Name=87;
CC Name=88;
CC Name=89;
CC Name=90;
CC Name=91;
CC Name=92;
CC Name=93;
CC Name=94;
CC Name=95;
CC Name=96;
CC Name=97;
CC Name=98;
CC Name=99;
CC Name=100;
CC Name=101;
CC Name=102;
CC Name=103;
CC Name=104;
CC Name=105;
CC Name=106;
CC Name=107;
CC Name=108;
CC Name=109;
CC Name=110;
CC Name=111;
CC Name=112;
CC Name=113;
CC Name=114;
CC Name=115;
CC Name=116;
CC Name=117;
CC Name=118;
CC Name=119;
CC Name=120;
CC Name=121;
CC Name=122;
CC Name=123;
CC Name=124;
CC Name=125;
CC Name=126;
CC Name=127;
CC Name=128;
CC Name=129;
CC Name=130;
CC Name=131;
CC Name=132;
CC Name=133;
CC Name=134;
CC Name=135;
CC Name=136;
CC Name=137;
CC Name=138;
CC Name=139;
CC Name=140;
CC Name=141;
CC Name=142;
CC Name=143;
CC Name=144;
CC Name=145;
CC Name=146;
CC Name=147;
CC Name=148;
CC Name=149;
CC Name=150;
CC Name=151;
CC Name=152;
CC Name=153;
CC Name=154;
CC Name=155;
CC Name=156;
CC Name=157;
CC Name=158;
CC Name=159;
CC Name=160;
CC Name=161;
CC Name=162;
CC Name=163;
CC Name=164;
CC Name=165;
CC Name=166;
CC Name=167;
CC Name=168;
CC Name=169;
CC Name=170;
CC Name=171;
CC Name=172;
CC Name=173;
CC Name=174;
CC Name=175;
CC Name=176;
CC Name=177;
CC Name=178;
CC Name=179;
CC Name=180;
CC Name=181;
CC Name=182;
CC Name=183;
CC Name=184;
CC Name=185;
CC Name=186;
CC Name=187;
CC Name=188;
CC Name=189;
CC Name=190;
CC Name=191;
CC Name=192;
CC Name=193;
CC Name=194;
CC Name=195;
CC Name=196;
CC Name=197;
CC Name=198;
CC Name=199;
CC Name=200;
CC Name=201;
CC Name=202;
CC Name=203;
CC Name=204;
CC Name=205;
CC Name=206;
CC Name=207;
CC Name=208;
CC Name=209;
CC Name=210;
CC Name=211;
CC Name=212;
CC Name=213;
CC Name=214;
CC Name=215;
CC Name=216;
CC Name=217;
CC Name=218;
CC Name=219;
CC Name=220;
CC Name=221;
CC Name=222;
CC Name=223;
CC Name=224;
CC Name=225;
CC Name=226;
CC Name=227;
CC Name=228;
CC Name=229;
CC Name=230;
CC Name=231;
CC Name=232;
CC Name=233;
CC Name=234;
CC Name=235;
CC Name=236;
CC Name=237;
CC Name=238;
CC Name=239;
CC Name=240;
CC Name=241;
CC Name=242;
CC Name=243;
CC Name=244;
CC Name=245;
CC Name=246;
CC Name=247;
CC Name=248;
CC Name=249;
CC Name=250;
CC Name=251;
CC Name=252;
CC Name=253;
CC Name=254;
CC Name=255;
CC Name=256;
CC Name=257;
CC Name=258;
CC Name=259;
CC Name=260;
CC Name=261;
CC Name=262;
CC Name=263;
CC Name=264;
CC Name=265;
CC Name=266;
CC Name=267;
CC Name=268;
CC Name=269;
CC Name=270;
CC Name=271;
CC Name=272;
CC Name=273;
CC Name=274;
CC Name=275;
CC Name=276;
CC Name=277;
CC Name=278;
CC Name=279;
CC Name=280;
CC Name=281;
CC Name=282;
CC Name=283;
CC Name=284;
CC Name=285;
CC Name=286;
CC Name=287;
CC Name=288;
CC Name=289;
CC Name=290;
CC Name=291;
CC Name=292;
CC Name=293;
CC Name=294;
CC Name=295;
CC Name=296;
CC Name=297;
CC Name=298;
CC Name=299;
CC Name=300;
CC Name=301;
CC Name=302;
CC Name=303;
CC Name=304;
CC Name=305;
CC Name=306;
CC Name=307;
CC Name=308;
CC Name=309;
CC Name=310;
CC Name=311;
CC Name=312;
CC Name=313;
CC Name=314;
CC Name=315;
CC Name=316;
CC Name=317;
CC Name=318;
CC Name=319;
CC Name=320;
CC Name=321;
CC Name=322;
CC Name=323;
CC Name=324;
CC Name=325;
CC Name=326;
CC Name=327;
CC Name=328;
CC Name=329;
CC Name=330;
CC Name=331;
CC Name=332;
CC Name=333;
CC Name=334;
CC Name=335;
CC Name=336;
CC Name=337;
CC Name=338;
CC Name=339;
CC Name=340;
CC Name=341;
CC Name=342;
CC Name=343;
CC Name=344;
CC Name=345;
CC Name=346;
CC Name=347;
CC Name=348;
CC Name=349;
CC Name=350;
CC Name=351;
CC Name=352;
CC Name=353;
CC Name=354;
CC Name=355;
CC Name=356;
CC Name=357;
CC Name=358;
CC Name=359;
CC Name=360;
CC Name=361;
CC Name=362;
CC Name=363;
CC Name=364;
CC Name=365;
CC Name=366;
CC Name=367;
CC Name=368;
CC Name=369;
CC Name=370;
CC Name=371;
CC Name=372;
CC Name=373;
CC Name=374;
CC Name=375;
CC Name=376;
CC Name=377;
CC Name=378;
CC Name=379;
CC Name=380;
CC Name=381;
CC Name=382;
CC Name=383;
CC Name=384;
CC Name=385;
CC Name=386;
CC Name=387;
CC Name=388;
CC Name=389;
CC Name=390;
CC Name=391;
CC Name=392;
CC Name=393;
CC Name=394;
CC Name=395;
CC Name=396;
CC Name=397;
CC Name=398;
CC Name=399;
CC Name=400;
CC Name=401;
CC Name=402;
CC Name=403;
CC Name=404;
CC Name=405;
CC Name=406;
CC Name=407;
CC Name=408;
CC Name=409;
CC Name=410;
CC Name=411;
CC Name=412;
CC Name=413;
CC Name=414;
CC Name=415;
CC Name=416;
CC Name=417;
CC Name=418;
CC Name=419;
CC Name=420;
CC Name=421;
CC Name=422;
CC Name=423;
CC Name=424;
CC Name=425;
CC Name=426;
CC Name=427;
CC Name=428;
CC Name=429;
CC Name=430;
CC Name=431;
CC Name=432;
CC Name=433;
CC Name=434;
CC Name=435;
CC Name=436;
CC Name=437;
CC Name=438;
CC Name=439;
CC Name=440;
CC Name=441;
CC Name=442;
CC Name=443;
CC Name=444;
CC Name=445;
CC Name=446;
CC Name=447;
CC Name=448;
CC Name=449;
CC Name=450;
CC Name=451;
CC Name=452;
CC Name=453;
CC Name=454;
CC Name=455;
CC Name=456;
CC Name=457;
CC Name=458;
CC Name=459;
CC Name=460;
CC Name=461;
CC Name=462;
CC Name=463;
CC Name=464;
CC Name=465;
CC Name=466;
CC Name=467;
CC Name=468;
CC Name=469;
CC Name=470;
CC Name=471;
CC Name=472;
CC Name=473;
CC Name=474;
CC Name=475;
CC Name=476;
CC Name=477;
CC Name=478;
CC Name=479;
CC Name=480;
CC Name=481;
CC Name=482;
CC Name=483;
CC Name=484;
CC Name=485;
CC Name=486;
CC Name=487;
CC Name=488;
CC Name=489;
CC Name=490;
CC Name=491;
CC Name=492;
CC Name=493;
CC Name=494;
CC Name=495;
CC Name=496;
CC Name=497;
CC Name=498;
CC Name=499;
CC Name=500;
CC Name=501;
CC Name=502;
CC Name=503;
CC Name=504;
CC Name=505;
CC Name=506;
CC Name=507;
CC Name=508;
CC Name=509;
CC Name=510;
CC Name=511;
CC Name=512;
CC Name=513;
CC Name=514;
CC Name=515;
CC Name=516;
CC Name=517;
CC Name=518;
CC Name=519;
CC Name=520;
CC Name=521;
CC Name=522;
CC Name=523;
CC Name=524;
CC Name=525;
CC Name=526;
CC Name=527;
CC Name=528;
CC Name=529;
CC Name=530;
CC Name=531;
CC Name=532;
CC Name=533;
CC Name=534;
CC Name=535;
CC Name=536;
CC Name=537;
CC Name=538;
CC Name=539;
CC Name=540;
CC Name=541;
CC Name=542;
CC Name=543;
CC Name=544;
CC Name=545;
CC Name=546;
CC Name=547;
CC Name=548;
CC Name=549;
CC Name=550;
CC Name=551;
CC Name=552;
CC Name=553;
CC Name=554;
CC Name=555;
CC Name=556;
CC Name=557;
CC Name=558;
CC Name=559;
CC Name=560;
CC Name=561;
CC Name=562;
CC Name=563;
CC Name=564;
CC Name=565;
CC Name=566;
CC Name=567;
CC Name=568;
CC Name=569;
CC Name=570;
CC Name=571;
CC Name=572;
CC Name=573;
CC Name=574;
CC Name=575;
CC Name=576;
CC Name=577;
CC Name=578;
CC Name=579;
CC Name=580;
CC Name=581;
CC Name=582;
CC Name=583;
CC Name=584;
CC Name=585;
CC Name=586;
CC Name=587;
CC Name=588;
CC Name=589;
CC Name=590;
CC Name=591;
CC Name=592;
CC Name=593;
CC Name=594;
CC Name=595;
CC Name=596;
CC Name=597;
CC Name=598;
CC Name=599;
CC Name=600;
CC Name=601;
CC Name=602;
CC Name=603;
CC Name=604;
CC Name=605;
CC Name=606;
CC Name=607;
CC Name=608;
CC Name=609;
CC Name=610;
CC Name=611;
CC Name=612;
CC Name=613;
CC Name=614;
CC Name=615;
CC Name=616;
CC Name=617;
CC Name=618;
CC Name=619;
CC Name=620;
CC Name=621;
CC Name=622;
CC Name=623;
CC Name=624;
CC Name=625;
CC Name=626;
CC Name=627;
CC Name=628;
CC Name=629;
CC Name=630;
CC Name=631;
CC Name=632;
CC Name=633;
CC Name=634;
CC Name=635;
CC Name=636;
CC Name=637;
CC Name=638;
CC Name=639;
CC Name=640;
CC Name=641;
CC Name=642;
CC Name=643;
CC Name=644;
CC Name=645;
CC Name=646;
CC Name=647;
CC Name=648;
CC Name=649;
CC Name=650;
CC Name=651;
CC Name=652;
CC Name=653;
CC Name=654;
CC Name=655;
CC Name=656;
CC Name=657;
CC Name=658;
CC Name=659;
CC Name=660;
CC Name=661;
CC Name=662;
CC Name=663;
CC Name=664;
CC Name=665;
CC Name=666;
CC Name=667;
CC Name=668;
CC Name=669;
CC Name=670;
CC Name=671;
CC Name=672;
CC Name=673;
CC Name=674;
CC Name=675;
CC Name=676;
CC Name=677;
CC Name=678;
CC Name=679;
CC Name=680;
CC Name=681;
CC Name=682;
CC Name=683;
CC Name=684;
CC Name=685;
CC Name=686;
CC Name=687;
CC Name=688;
CC Name=689;
CC Name=690;
CC Name=691;
CC Name=692;
CC Name=693;
CC Name=694;
CC Name=695;
CC Name=696;
CC Name=697;
CC Name=698;
CC Name=699;
CC Name=700;
CC Name=701;
CC Name=702;
CC Name=703;
CC Name=704;
CC Name=705;
CC Name=706;
CC Name=707;
CC Name=708;
CC Name=709;
CC Name=710;
CC Name=711;
CC Name=712;
CC Name=713;
CC Name=714;
CC Name=715;
CC Name=716;
CC Name=717;
CC Name=718;
CC Name=719;
CC Name=720;
CC Name=721;
CC Name=722;
CC Name=723;
CC Name=724;
CC Name=725;
CC Name=726;
CC Name=727;
CC Name=728;
CC Name=729;
CC Name=730;
CC Name=731;
CC Name=732;
CC Name=733;
CC Name=734;
CC Name=735;
CC Name=736;
CC Name=737;
CC Name=738;
CC Name=739;
CC Name=740;
CC Name=741;
CC Name=742;
CC Name=743;
CC Name=744;
CC Name=745;
CC Name=746;
CC Name=747;
CC Name=748;
CC Name=749;
CC Name=750;
CC Name=751;
CC Name=752;
CC Name=753;
CC Name=754;
CC Name=755;
CC Name=756;
CC Name=757;
CC Name=758;
CC Name=759;
CC Name=760;
CC Name=761;
CC Name=762;
CC Name=763;
CC Name=764;
CC Name=765;
CC Name=766;
CC Name=767;
CC Name=768;
CC Name=769;
CC Name=770;
CC Name=771;
CC Name=772;
CC Name=773;
CC Name=774;
CC Name=775;
CC Name=776;
CC Name=777;
CC Name=778;
CC Name=779;
CC Name=780;
CC Name=781;
CC Name=782;
CC Name=783;
CC Name=784;
CC Name=785;
CC Name=786;
CC Name=787;
CC Name=788;
CC Name=789;
CC Name=790;
CC Name=791;
CC Name=792;
CC Name=793;
CC Name=794;
CC Name=795;
CC Name=796;
CC Name=797;
CC Name=798;
CC Name=799;
CC Name=800;
CC Name=801;
CC Name=802;
CC Name=803;
CC Name=804;
CC Name=805;
CC Name=806;
CC Name=807;
CC Name=808;
CC Name=809;
CC Name=810;
CC Name=811;
CC Name=812;
CC Name=813;
CC Name=814;
CC Name=815;
CC Name=816;
CC Name=817;
CC Name=818;
CC Name=819;
CC Name=820;
CC Name=821;
CC Name=822;
CC Name=823;
CC Name=824;
CC Name=825;
CC Name=826;
CC Name=827;
CC Name=828;
CC Name=829;
CC Name=830;
CC Name=831;
CC Name=832;
CC Name=833;
CC Name=834;
CC Name=835;
CC Name=836;
CC Name=837;
CC Name=838;
CC Name=839;
CC Name=840;
CC Name=841;
CC Name=842;
CC Name=843;
CC Name=844;
CC Name=845;
CC Name=846;
CC Name=847;
CC Name=848;
CC Name=849;
CC Name=850;
CC Name=851;
CC Name=852;
CC Name=853;
CC Name=854;
CC Name=855;
CC Name=856;
CC Name=857;
CC Name=858;
CC Name=859;
CC Name=860;
CC Name=861;
CC Name=862;
CC Name=863;
CC Name=864;
CC Name=865;
CC Name=866;
CC Name=867;
CC Name=868;
CC Name=869;
CC Name=870;
CC Name=871;
CC Name=872;
CC Name=873;
CC Name=874;
CC Name=875;
CC Name=876;
CC Name=877;
CC Name=878;
CC Name=879;
CC Name=880;
CC Name=881;
CC Name=882;
CC Name=883;
CC Name=884;
CC Name=885;
CC Name=886;
CC Name=887;
CC Name=888;
CC Name=889;
CC Name=890;
CC Name=891;
CC Name=892;
CC Name=893;
CC Name=894;
CC Name=895;
CC Name=896;
CC Name=897;
CC Name=898;
CC Name=899;
CC Name=900;
CC Name=901;
CC Name=902;
CC Name=903;
CC Name=904;
CC Name=905;
CC Name=906;
CC Name=907;
CC Name=908;
CC Name=909;
CC Name=910;
CC Name=911;
CC Name=912;
CC Name=913;
CC Name=914;
CC Name=915;
CC Name=916;
CC Name=917;
CC Name=918;
CC Name=919;
CC Name=920;
CC Name=921;
CC Name=922;
CC Name=923;
CC Name=924;
CC Name=925;
CC Name=926;
CC Name=927;
CC Name=928;
CC Name=929;
CC Name=930;
CC Name=931;
CC Name=932;
CC Name=933;
CC Name=934;
CC Name=935;
CC Name=936;
CC Name=937;
CC Name=938;
CC Name=939;
CC Name=940;
CC Name=941;
CC Name=942;
CC Name=943;
CC Name=944;
CC Name=945;
CC Name=946;
CC Name=947;
CC Name=948;
CC Name=949;
CC Name=950;
CC Name=951;
CC Name=952;
CC Name=953;
CC Name=954;
CC Name=955;
CC Name=956;
CC Name=957;
CC Name=958;
CC Name=959;
CC Name=960;
CC Name=961;
CC Name=962;
CC Name=963;
CC Name=964;
CC Name=965;
CC Name=966;
CC Name=967;
CC Name=968;
CC Name=969;
CC Name=970;
CC Name=971;
CC Name=972;
CC Name=973;
CC Name=974;
CC Name=975;
CC Name=976;
CC Name=977;
CC Name=978;
CC Name=979;
CC Name=980;
CC Name=981;
CC Name=982;
CC Name=983;
CC Name=984;
CC Name=985;
CC Name=986;
CC Name=987;
CC Name=988;
CC Name=989;
CC Name=990;
CC Name=991;
CC Name=992;
CC Name=993;
CC Name=994;
CC Name=995;
CC Name=996;
CC Name=997;
CC Name=998;
CC Name=999;
CC Name=1000;
CC Name=1001;
CC Name=1002;
CC Name=1003;
CC Name=1004;
CC Name=1005;
CC Name=1006;
CC Name=1007;
CC Name=1008;
CC Name=1009;
CC Name=1010;
CC Name=1011;
CC Name=1012;
CC Name=1013;
CC Name=1014;
CC Name=1015;
CC Name=1016;
CC Name=1017;
CC Name=1018;
CC Name=1019;
CC Name=1020;
CC Name=1021;
CC Name=1022;
CC Name=1023;
CC Name=1024;
CC Name=1025;
CC Name=1026;
CC Name=1027;
CC Name=1028;
CC Name=1029;
CC Name=1030;
CC Name=1031;
CC Name=1032;
CC Name=1033;
CC Name=1034;
CC Name=1035;
CC Name=1036;
CC Name=1037;
CC Name=1038;
CC Name=1039;
CC Name=1040;
CC Name=1041;
CC Name=1042;
CC Name=1043;
CC Name=1044;
CC Name=1045;
CC Name=1046;
CC Name=1047;
CC Name=1048;
CC Name=1049;
CC Name=1050;
CC Name=1051;
CC Name=1052;
CC Name=1053;
CC Name=1054;
CC Name=1055;
CC Name=1056;
CC Name=1057;
CC Name=1058;
CC Name=1059;
CC Name=1060;
CC Name=1061;
CC Name=1062;
CC Name=1063;
CC Name=1064;
CC Name=1065;
CC Name=1066;
CC Name=1067;
CC Name=1068;
CC Name=1069;
CC Name=1070;
CC Name=1071;
CC Name=1072;
CC Name=1073;
CC Name=1074;
CC Name=1075;
CC Name=1076;
CC Name=1077;
CC Name=1078;
CC Name=1079;
CC Name=1080;
CC Name=1081;
CC Name=1082;
CC Name=1083;
CC Name=1084;
CC Name=1085;
CC Name=1086;
CC Name=1087;
CC Name=1088;
CC Name=1089;
CC Name=1090;
CC Name=1091;
CC Name=1092;
CC Name=1093;
CC Name=1094;
CC Name=1095;
CC Name=1096;
CC Name=1097;
CC Name=1098;
CC Name=1099;
CC Name=1100;
CC Name=1101;
CC Name=1102;
CC Name=1103;
CC Name=1104;
CC Name=1105;
CC Name=1106;
CC Name=1107;
CC Name=1108;
CC Name=1109;
CC Name=1110;
CC Name=1111;
CC Name=1112;
CC Name=1113;
CC Name=1114;
CC Name=1115;
CC Name=1116;
CC Name=1117;
CC Name=1118;
CC Name=1119;
CC Name=1120;
CC Name=1121;
CC Name=1122;
CC Name=1123;
CC Name=1124;
CC Name=1125;
CC Name=1126;
CC Name=1127;
CC Name=1128;
CC Name=1129;
CC Name=1130;
CC Name=1131;
CC Name=1132;
CC Name=1133;
CC Name=1134;
CC Name=1135;
CC Name=1136;
CC Name=1137;
CC Name=1138;
CC Name=1139;
CC Name=1140;
CC Name=1141;
CC Name=1142;
CC Name=1143;
CC Name=1144;
CC Name=1145;
CC Name=1146;
CC Name=1147;
CC Name=1148;
CC Name=1149;
CC Name=1150;
CC Name=1151;
CC Name=1152;
CC Name=1153;
CC Name=1154;
CC Name=1155;
CC Name=1156;
CC Name=1157;
CC Name=1158;
CC Name=1159;
CC Name=1160;
CC Name=1161;
CC Name=1162;
CC Name=1163;
CC Name=1164;
CC Name=1165;
CC Name=1166;
CC Name=1167;
CC Name=1168;
CC Name=1169;
CC Name=1170;
CC Name=1171;
CC Name=1172;
CC Name=1173;
CC Name=1174;
CC Name=1175;
CC Name=1176;
CC Name=1177;
CC Name=1178;
CC Name=1179;
CC Name=1180;
CC Name=1181;
CC Name=1182;
CC Name=1183;
CC Name=1184;
CC Name=1185;
CC Name=1186;
CC Name=1187;
CC Name=1188;
CC Name=1189;
CC Name=1190;
CC Name=1191;
CC Name=1192;
CC Name=1193;
CC Name=1194;
CC Name=1195;
CC Name=1196;
CC Name=1197;
CC Name=1198;
CC Name=1199;
CC Name=1200;
CC Name=1201;
CC Name=1202;
CC Name=1203;
CC Name=1204;
CC Name=1205;
CC Name=1206;
CC Name=1207;
CC Name=1208;
CC Name=1209;
CC Name=1210;
CC Name=1211;
CC Name=1212;
CC Name=1213;
CC Name=1214;
CC Name=1215;
CC Name=1216;
CC Name=1217;
CC Name=1218;
CC Name=1219;
CC Name=1220;
CC Name=1221;
CC Name=1222;
CC Name=1223;
CC Name=1224;
CC Name=1225;
CC Name=1226;
CC Name=1227
```

DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
 DR GO; GO:0016080; P:synaptic vesicle targeting; NAS.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR008899; Znf_piccolo.
 DR Pfam; PF00168; C2; 2.
 DR Pfam; PF00595; PDZ; 1.
 DR Pfam; PF05715; Zf_piccolo; 2.
 DR SMART; SM00239; C2; 2.
 DR SMART; SM00228; PDZ; 1.
 DR PROSITE; PS00499; C2_DOMAIN_1; 1.
 DR PROSITE; PS00004; C2_DOMAIN_2; 2.
 DR PROSITE; PS0106; PDZ; 1.
 KW Calcium/phospholipid-binding; Metal-binding; Zinc; zinc-finger;
 Repeat; Alternative splicing.
 FT DOMAIN 372 491
 FT ZN_FING 523 547
 FT ZN_FING 1010 1033
 FT DOMAIN 2351 2362
 FT DOMAIN 4442 4536
 FT DOMAIN 4653 4752
 FT DOMAIN 4968 5059
 FT VARSPIC 4876 4880
 FT VARSPIC 4881 5085
 FT MUTAGEN 4668 4668
 FT MUTAGEN 4674 4674
 FT MUTAGEN 4688 4688
 FT MUTAGEN 4688 4689
 FT MUTAGEN 4689 4689
 FT MUTAGEN 4690 4691
 FT MUTAGEN 4692 4693
 FT MUTAGEN 4694 4694
 FT SEQUENCE 5085 AA; 552702 MW; 5A1BB543201A7450 CRC64;
 Query Match 3.9%; Score 423; DB 1; Length 5085;
 Best local Similarity 19.9%; Pred. No. 6,3e-07;
 Matches 404; Conservative 270; Mismatches 757; Indels 596; Gaps 87;
 21 PAPAPNKGSSSSVGEKLGAAADA VGRTEYRRRRHTMDKDSGAAT--TTTHER 78
 1087 PMPAPAPESPQQPTGTQVKGKKKEA--EGKTE--AEKPEKETASIEKTPMTTDDK 1142
 79 FRRRSVTCNSMATALPGLPLSLP-----OPSI-----PAVPOS 114
 1143 -----LESEEGKKSYSALPEKKPEEEKKASADKERRKPAEEKRPLEKKPIVDCK 1196
 115 APPPEP-----RETVTATSQVAQPP--AAAPGE-----QAVAPAPSTV 156
 1197 LPPEAKPLSSEGEKKEIKKAVQIPEEPTGKVAAKAGEEQQDPSREALPGATPLTL 1256
 1257 PRAGEKERAVAPQAGSSKDDGEGHS-----KEKTEKEDKSDTSQQPKSPQGIS 1309
 157 PESTSKDRPVSQPSILVSGKEPPPARSGSGSAGKEPQERSQQDDIDELETKA----- 211
 157 PRAGERAVAPQAGSSKDDGEGHS-----KEKTEKEDKSDTSQQPKSPQGIS 1309
 212 -VGSNDGFLKFDLIEIGSGFTYVKGLDTETVEVANCEDQRKLTSEORFEEBAE 270
 1310 DTGYSDDG-----ISGLGEIPLSLP-----SDEKD 1335
 271 MKGGOHPPIVAFYSWEST-----VKGKCIIVLVELMTSGTL-KTYLKRFPYMKIK 322
 1336 LKGLKKDSF-----SQESSPSPSDLALKESTVSLIEAQA STLVGKAEKKTPOKIS 1390

QY 323 VLRSKQILKGLQFLATRPPIIHRDKCDNPIFGTSGVKIGDLATLKRAFPKAS 382
 Db 1391 PEKQDDQ-----KTQT-----ASPTLDTITSEEIKSQEKSPK 1427
 QY VIGTPEPAPMEYKEDSDVYVAFGCMLEMTSEVPY-----SECNAQIYRYVT 436
 Db 1428 KDSQGPSPKREHK--PELVD-----DLSPRASYSVEDSESESPVRRRRR 1477
 QY 437 SGVPASFDKVAIPVEVEILIEGCIKONKDERYSKDLINHPQDET-----VRVLA 490
 Db 1478 TSIGSSSD-----EYQ--EDSGSGEEDPIKQIIEKMADEBAGSEDEEFTSOLK 1530
 QY 491 E-EDDGKAIKMLRIEDIKLKKYK-----DN 519
 Db 1531 EISGVGES-----QKREBAKGGKGVAGKRRILTRKSTSPDDAGRRHSHDEDD 1581
 QY 520 EALFEFDLE-RDVPEDVAQEMVSGYVCEGDHT--MAAKIKRVSLIKRKRQRLVR 576
 Db 1582 ETPDESEFLKFRETKSGSESELVVAAGGGLRRFTIENSTIADKYS-----SESSQKKTIL 1638
 QY 577 EDEKKEKQESSSLKQVQESSASQGTGKOLPSAS--TGIPATSTASVSTQVEPEEPA 634
 Db 1639 YPDEPEPELMEBSLDPEDSRGE--GSSLHASFPTGTPTSVSLDEDSDESPSHKG 1697
 QY 635 DQHQQLQYQPPSISVSDGTVDGSGSSVFTESRVSSQQTVSYGSGHQHASTGTVDGHI 694
 Db 1698 ESQQRKARHRSHGPIPTIEDSESEELKEEELKEQKQRELEQQRKSSKSKKD 1757
 QY 695 PSTVQAG-SQPHGVPPSSVAQGSQGSQSSSLTGVSQPIQHPPQGGIQTAPQ 753
 Db 1758 KDEIRKQRRERKPTPSNLS-----PIEDASPTELQQAEMEE 1797
 QY 754 TVQYSLSQTSSTSEATTAQVPSQAPVLPQVSAQKSTQGSQVAPAPVVAQPAQ 813
 Db 1798 IHRSSGSEVSPSTE--SDPEGEISPEKILIEQYKILPAVSLYSFPTDQGVQKEGV 1854
 QY 814 QPTLLASSVDASDVASGMSDGENAVPSSGHEGRTTTPHYRKSVRSRHEKTSRPK 873
 Db 1855 Q-----KALKSAEYEEEMQKPH-----K 1874
 QY 874 LRILANSKQDRVVECO-----LETHNRKXV--TEKFDLGD--NPEELATIMVN 920
 Db 1875 YKAFPAANRDEVEFEPELYGGMILIEDYIESVEDTYNGSVGSLTROEE-----QN 1928
 QY 921 DFLALIERSPVDQVREILEKADEMLSDEVYVEPDQGLSELOGKDPYFS----- 972
 Db 1929 GFMOQGRGQKVRLOQIYDDPMOKIS-DLOKEFELESLHSVPPQEDIVSSYIIPESH 1987
 QY 973 -----GSQKLE-GEFKQPIPAS-----MPOQIGLPTSSLQGVHSAGRFTVSPVES 1020
 Db 1988 EIVDLSGMYMSTSEBKLDADSAVEELMRQOVYTDGS-----SPQVTIGDDMAES 2041
 QY 1021 RRESKVPFSEITDPAVASTAOSPQNLSHSASLSIQ-----QAFS--ELRQKMT 1071
 Db 2042 TLDPDRVQDASLTSLS-----QASLTDSYSSATLSIPDKYITQFAEBLEDEYVD 2095
 QY 1072 GNTAPNPNESHGTPFPVPPFLSIAVPTTAAATAPVAPNAPSPPNDIS--TSVQSEV 1129
 Db 2096 YTREIODITLHBSLILTYSEPEES-----ATSVPSDDIPSLTSSISVC 2139
 QY 1130 TVPTREGIAGVATSGVVTSGGLPIPVSESPLVSVV--SGITPAVVIS--TSP 1183
 Db 2140 TTDSSSPYTTLDLTLVYTEPADVMTKFDSEHISSTYPGSIIDIPEDISVLDRTIMP 2199
 QY 1184 SLQVPTSTSEIYVSTALYPSVTVSATSSAGSGATATGPKRPVAVSQAGASTVIG-- 1240
 Db 2200 BSR--TNEDRIYVLSGVAIPSVESVGTRPPEQADTISTDPISEKDLIKQKKEGTGCI 2257
 QY 1241 -----ATLTSVSTTSPSTASQSLQSLSSSTPTLAETVVAHSL-- 1283
 Db 2258 ILEVLDAYKDRKESEALTKISLPE--PGA-----QAPSVVTPQIKQ--HVSPPHVS 2309

```

QY 1284 -----DKTSHSSTGIALFSLAPSSSSPGAGVSS--YISQPGIHLPIVPSVLA--- 1332
DB 2310 GKISGGEKTYRLPSG-SLPVSTHPKSRPFRRSSLDLSADPPPPPPSPSSSP 2368
QY 1333 -TPILPQAGP-----TSTPLIPQVPSIPPLVQVPA-----NPAVQOT 1370
DB 2369 PFPPLPAPSPKRPYPPKKKLAVALVTSTIYVTHVDALTWEEAAAARRSNGLPATKMC 2428
QY 1371 LHS---QOPALIP-NQPHITCPV-----DSTQKAPSIDDI 1406
DB 2429 ALAPPVPKPKPSQIPGLVFTHRPEAIKPIAPKPAVPOI PVTQKPTCTCPKPTLSLT 2488
QY 1407 KTLREEL-----SLFSEHSSSGAQAHSVSL-ETSLVIESVTPGIPPTVA-- 1452
DB 2489 STWSLNLVSADYNVSPSPSPSPKSSPRKSKLMOTVITLPSFGPTDSSAAQ 2548
QY 1453 -----PSKLLTS-----TSTCLPPTNLP-----LGTVALPVTPTV-- 1484
DB 2549 AITSMPLGSPKOLVLETFVSVPPMTSTETPSAQPTLYTSALGATPS--VTPAVTAS 2606
QY 1485 -----PGQVSTPVSTTSGVKGTAPSK--PLTKAPV-----LP 1517
DB 2607 LFGVPTSLTQFLPAESKKEVASVAVSVAPRSVSIPIPEPLALDRHOKENGKLP 2666
QY 1518 -VGTSLPAGTLPEEQLPPFPGSLTQSQPLE-----DLDAQ-----LRRTLS--PEXI 1563
DB 2667 LIGDAIDLRTIPKSEV-----KYTECMPLSAMDVKRQTLANVEYRRQISAVGPSII 2720
QY 1564 TVTSAVG---PVSM-----AAPTAITEAGTPQKGVSVQVKEGVLATSSGAGVFKM 1611
DB 2721 NLSAASLGLPVTMDSKTVAVVTCTDTIYTTGTSGVQIEHHTVSLQLTTS----- 2773
QY 1612 GREGVSAAGAQKEGKNEKEDAKSVHFESSTSESSVSSSPSS--TLVKDEP----- 1663
DB 2774 -----KHTLEPYRKPSQAFTIRDEAPINLSLGSQAQVTLATVTKPVTVP 2820
QY 1664 -----NGITIPGISDVPE-----SAKTTASEKSDTGOPTKV 1697
DB 2821 VGVTNGMTDSTLSQVADGEVLDLSTKSHRTVVTMDSEISNVYTKI 2867

```

```

RESULT 11
NRK1_YEAST
ID NRK1_YEAST STANDARD; PRT: 1080 AA.
AC P38692;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase NRK1 (EC 2.7.1.37) (N-rich kinase 1).
GN NRK1 OR KIC1 OR YHR102W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC-5;
RA Fukami Y.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SG288C / AB972;
RC MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Gelsel C., Kisten J.,
RA Kucera T., Hillier L.W., Jier M., Johnston L., Langston Y.,
RA Lareille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nham M., Rifkin L., Riles L., St Peter H., Trevisan S., Vaughan K.,
RA Vignati D., Wilcox L., Woldman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL Science 265:2077-2082 (1994).

```

```

CC -! CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -! SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: D29880; BA06250.1; -.
CC EMBL: U00059; AB06860.1; -.
CC PIR: S48944; S48944.
CC Germonline: 139419; -.
CC SGD: S0001144; KIC1.
CC GO: GO:0016301; P-kinase activity; IDR.
CC GO: GO:0007047; P-cell wall organization and biogenesis; IGI.
CC GO: GO:0009092; P-cellular morphogenesis; IGI.
CC InterPro: IPR000719; Prot kinase.
CC InterPro: IPR008271; Ser Thr pkin AS.
CC InterPro: IPR002290; Ser thr kinase.
CC InterPro: IPR001245; Tyr_kinase.
CC Pfam: PFO0069; Pkinase; 1.
CC PRINTS: PR00109; TYRKINASE.
CC ProDom: PD000001; Prot kinase; 1.
CC SMART: SM00220; S_TKc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
CC TRANSFERASE: Serine/threonine-protein kinase; ATP-binding.
CC DOMAIN 23 276
CC FT NP_BIND 29 37 ATP (BY SIMILARITY).
CC FT BINDING 52 52 ATP (BY SIMILARITY).
CC FT ACT_SITE 144 144 BY SIMILARITY.
CC SQ SEQUENCE 1080 AA; 117061 MW; 99899EAF315E0B94 CRC64;

```

```

Query Match 3.8%; Score 416; DB 1; Length 1080;
Best Local Similarity 22.1%; Pred. No. 2.1e-07;
Matches 257; Conservative 173; Mismatches 473; Indels 262; Gaps 52;

```

```

QY 227 IGRGSEFTYKGLDTEITVEVANCELODRKLTSEKQFEEBEMLKG-LQHNIYRFYD 285
DB 29 IGRGSEFTYKGLDTEITVEVANCELODRKLTSEKQFEEBEMLKG-LQHNIYRFYD 87
QY 286 SMESTYKGGKCIYLVNELMNGSLTKYKFKFKYKIKVLSMCRQILKGHFLHTPTPI 345
DB 88 SYLKDTLS---LWIMEHCAGSLRSILRFGKIDE-KYIGVIRBELVALKCIHKON--V 140
QY 346 IHRDLKCDNIFITGPTSVKIGDLGLA-----TLKRASPAKSVITGTFEPMADPMTEE-- 397
DB 141 IHRDLKCDNIFITGPTSVKIGDLGLA-----TLKRASPAKSVITGTFEPMADPMTEE-- 195
QY 398 KYDESVDVYAFGCMLEMAISEYPSYSECQMAQIKRYRVTSGVVPASDKAIVEKELIE 457
DB 196 KYDESVDVYAFGCMLEMAISEYPSYSECQMAQIKRYRVTSGVVPASDKAIVEKELIE 254
QY 458 GCIKRONDERYSIKDLINHAFFQETGVRVLAEDDEGEKIAIKLWLRIDIKLKQKYY 517
DB 255 LCLDEDEKERSADDLKSKFIRAKHKTPTPSILKE-----LISRYLFPD--KKNKKYK 306
QY 518 DNEAIEBSPDLERDVPEDVAQEVESGVYCE--GDHKYAKAIKDRVSLIKRREBQRL 574
DB 307 -----IGSIPENBPSKPEAPKPSQNGSGD-----EAKKSI 338
QY 575 VREQEKKKQDESSLKQOVESKASQIGI-----KQIPASATGIPASTTSASVSTQVER 629
DB 339 VREQEKKKQDESSLKQOVESKASQIGI-----KQIPASATGIPASTTSASVSTQVER 676
QY 630 E-----PEADQHQQLQYQOPSSISVLDGTVDSGGQSSV-----FTSRVSSQQTQS 676
DB 399 DEDSYDPTSHNTRPFTVYQCTTIIGKVGPGII--AQNSTLNAVPTNNYTNKSPKQVAG 456

```

677 YGSGHEQAHSTGT-VPCHIPSTVQAQSQ-----PQGVYPP 710
 Db 457 -----TNTSGTHAGPWTSSKRLSKAPKQLLEFENELITAMENDVNTAPKISISI 510
 QY 711 SSVAGGQGOQ---PS-SSSLTG-VSSGQPIQH-PQOQOQ---IQOTAPPOQTVQVLSLQ 761
 Db 511 SSVNAGNSRRDPFIPSTISNEVNGNINNNKMRPPLSSGNNVYQSQTALPFLITQTKFNK 570
 QY 762 TS---TSSEATTAQVPOAPQAPVPOVYAGKOSTQGVASQVAPAPVAVQAQAPQPTL 818
 Db 571 TSKGPTTGLTAPITAPISLEIPELP-----NSALPTPASADPVLIT-----PSTK 615
 QY 819 ASSVDSANSDVYASGMDGNEVPSSSGRHEGRTTKRHYKSVRSRHE-----KTSRPL 874
 Db 616 ARS-----STVYAG-----TPSSSSSIQYKSPENVRRILTVSNRRHEHCSTITNQKL 663
 QY 875 RILVNSNGKGRVVEQQLTHNRKQVTFPFDLDGNDPEIATITMNDITALEBSFVDQ 934
 Db 664 GSAVASNGISTPNSNNNN-----NTDSENSRSGSSSNTANSTOMGITPQVNTK 716
 QY 935 VREIIEKADEMLSDVSVPEPDQGLBSLQCKDDYGFSGQKLEGEFPKQPIPASSMPQOI 994
 Db 717 -----LSTHKASSPS-----RPLFG---VGTSPNK-----PAASPTONI 748
 QY 995 GIPTSSTLVVHSGKRFIVSPVESRLRESK---VFPSEITDTVAASTQSPGMILSH 1050
 Db 749 G-----HNSHTNLAPPTMKPMANSKXNDKILLQPLNSIPSSSTLNTISGN--- 795
 QY 1051 SASLSLQQAASELRRQMTEGPNTPAPNPHSHOPTFPVPPFLSSLAGVTTAAATAPV 1110
 Db 796 SSNNLTSSNFSNKESSRRNG-----DEKRNPNLKLQWPL-----PT 834
 QY 1111 PATSSPNDISTSVIOGEVTPTEGIAVATST-----GVWTSGLPIPPVSPPLSS 1165
 Db 835 PVVANKLLDPNTASQNNNGP---GSAGITNENINQFENISSASNI-PVSNTPISEK 890
 QY 1166 VVS-SITTPAVVSTSTSPSLQVPTSTSEIVVSTALPVSVVATNSASAGSTATPGRPK 1224
 Db 891 HIFPGKIKXSQSTSNRKN-----SASEHPNLITLSSVSGVNGVGINNNVSGNNGN 945
 QY 1225 PPAVSSQAAGSTTVGATLTSVSTTTSFPSTASQLSTQLSSSTSTPLATVVVA---H 1281
 Db 946 NSVPLSAN-TGNTTIRKANTTITANSS-----AAASTTAPISQOTIPSGTQFNH 993
 QY 1282 SLDK--TSHSSTGLAFSLAPSSS 1304
 Db 994 ILSAATPAANSVNSIGCMCPPOS 1018

RESULT 12
 NC06_MOUSE STANDARD; PRT; 2067 AA.
 AC Q9JLI9; Q9JLI9;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Nuclear receptor coactivator 6 (Amplified in breast cancer-3 protein)
 DE signal cointegrator-2 (ASC-2) (Peroxisome proliferator-activated
 DE receptor-interacting protein) (PPAR-interacting protein) (Nuclear
 DE receptor-activating protein, 250 kDa) (Nuclear receptor coactivator
 DE RAP250) (NRC) (Thyroid hormone receptor binding protein).
 GN NC046 OR AIB3 OR RAP250 OR PRIP OR TRBP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1 AND 2), AND INTERACTION WITH
 RP PPARA; PPARA; RARA; RXRA; ESR1; ESR2 AND THR8.
 RC TISSUE=Liver;
 RX MEDLINE=20250907; PubMed=1078465;
 RA Zhu Y.-J., Kan L., Qi C., Kanwar Y.S., Yeldandi A.V., Rao M.S.,

RA Reddy J.K.;
 RT "Isolation and characterization of peroxisome proliferator-activated
 RT receptor (PPAR) interacting protein (PRIP) as a coactivator for
 RT PPAR.";
 RL J. Biol. Chem. 275:13510-13516(2000).
 RN [2]
 RP TISSUE=FROM N.A. (ISOFORM 2).
 RC TISSUE=Breast;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Faley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 786-1142 FROM N.A. (ISOFORM 1), INTERACTION WITH PPARA;
 RP PPARA; ESR1; ESR2; THRA AND THR8, AND MUTAGENESIS OF LEU-891 AND
 RP LEU-894.
 RC TISSUE=Embryo;
 RX MEDLINE=20148724; PubMed=10681503;
 RA Gaira F., Antonson P., Pelto-Huikko M., Treutner E., Gustafson J.-A.;
 RT "Cloning and characterization of RAP250, a nuclear receptor
 RT coactivator.";
 RL J. Biol. Chem. 275:5308-5317(2000).
 RN [4]
 RP INTERACTION WITH RNP2.
 RX MEDLINE=21638469; PubMed=11704680;
 RA Jung D.-J., Na S.-Y., Na D.S., Lee J.W.;
 RT "Molecular cloning and characterization of CAPER, a novel coactivator
 RT of activating protein-1 and estrogen receptors.";
 RL J. Biol. Chem. 277:1229-1234(2002).
 RC -
 CC FUNCTION: Nuclear receptor coactivator that directly binds nuclear
 CC receptors and stimulates the transcriptional activities in a
 CC hormone-dependent manner. Coactivates expression in an agonist-
 CC and AF2-dependent manner. Involved in the coactivation of
 CC different nuclear receptors, such as for steroids (GR and ERs),
 CC retinoids (RARs and RXRs), thyroid hormone (TRs), vitamin D3 (VDR)
 CC and prostanoids (PPARs). Probably functions as a general
 CC coactivator, rather than just a nuclear receptor coactivator.
 CC May coactivate expression via a remodeling of the NF-kappa-B pathway.
 CC interaction with histone acetyltransferase proteins. Involved in
 CC Placental, cardiac, hepatic and embryonic development.
 CC -
 CC SUBUNIT: Monomer and homodimer. Interacts in vitro with the basal
 CC transcription factors GTF2A and TBP, suggesting an autonomous
 CC transactivation function. Interacts with NC0A1, CREB3, RBM14, the
 CC histone acetyltransferase proteins EP300 and CREBBP, and with
 CC methyltransferase proteins NC0A6 and HMTYLI (By similarity).
 CC Interacts with RNP2. Belongs to the ASC-2/NC0A6 complex (ASCOM),
 CC which contains ASC-2/NC0A6, the retinoblastoma-binding protein
 CC RBQ-3/RBBP5, alpha- and beta-tubulins, the trithorax group
 CC proteins MLL2 and MLL3, and ASH2/ASCL2 (By similarity).
 CC -
 CC SUBCELLULAR LOCATION: Nuclear.
 CC -
 CC ALTERNATIVE PRODUCTS: Nuclear.
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9JLI9-1; Sequence=Displayed;
 CC Name=2;

```

QY 569 KEORQVAREOEKKKKKESSLKQVVEBSSAQT----- 601
Db 261 ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 602 GIKOLPSASTGPIPTASTTSASVSTOVEEPEEPADHQOOLQYO-OFESIVSLSD----- 652
Db 321 GWNQPSGALOPPPAOGSLGTHTTQMGKKAPLPSPMQAQLQARPSLATVQOTPSHPPPY 360
QY 653 --GTVDSQGGSSVFTESPVSQOQTVSY--GSGH-----EQAHSTGTVPCHIPS-- 696
Db 381 PFGSQASQAHNPFQWPNBQGFAPQMKGLOGGGSRVPTPLQOPHLITKSPASSPSFQ 440
QY 697 -----TVQAQSPGHVYPPSPSSVAQQGSGQSPSSSLTGVSSSQPIQHQQQOG---- 744
Db 441 QGSPASSPTVQOTQOQMGPRPPQNNPLSQGFQOQPVSS-----FGKNPWQQGNVPP 491
QY 745 ---IQOTAPPOQTVQ-----YLSLQSTSTS-----SEATTAPVSOQAP 780
Db 492 NFMWVQQQPPNGPQPSLHGLGLGMPKRLPFGFSAGANFNMGGQVPTTAATPGNSGAL 551
QY 791 QVLPQVSAKGKOSTQGVQVAPAEFVAVAQOPQATOPTLLASSVDSASDVASGMSDQENY 840
Db 552 QLOANQNVQAHAGQGAQGPQOMQOVSHGPPNMQF-----SLMGTHGNI--NNQQAQSSGV 605
QY 841 PS-SSGRHGRRTT-----RHYRQSVSRSSHEKTSRPK---IRLLVNSKGR 885

```

[illegible]

QY 110 AVPOSAPREHREYATATATSOVAQAQPPAAAPGQAVAGAPASTVPSSTKDRVPSOP 169
 Db 910 SQPVSKRPATPOKQPPAVSHS-----PQKSTPTPATKREK 950
 QY 170 SLVSGKEPPPARSSGGGSAKEPQERSQOQODILEETKAVGNSNDKFLKFDIEIKR 229
 Db 951 SV---PKEVPKLOQG-----KLEKTLSDAKLOQGIQKEDAKSKOKKLPK-----T 992
 QY 230 GSEKTVYKLDIETTVAVAMCELODKRLTKSERQKREAEMLKGIQHNIYRFYDWS 289
 Db 993 PSADKIQRVOKEDS-----RLQOTKLT-----PSSDKILHGVOKEDI----- 1032
 QY 290 TVAGKCIYLVTELMSTGLTKYLKPKYMKIKVLKSWCRQILKGLFLHTRTPPIIRD 349
 Db 1033 -----KFOAKLAKIS-----ADKLHL 1052
 QY 350 LKDNIFITPTGSVKIGLGLATLKRASFASVIGTPEFMADEMYEKYDESVDVYAFG 409
 Db 1053 QKED-----PKLOQMKAKAL---SADKIQPEAQKE----- 1080
 QY 410 MCMLEMATSEYPSSECONAAQIYRRTSGVKPASFPKVAIPYKELIEGIRQKDERYS 469
 Db 1081 ----DVOLOEVLRSKAVSADKIQHGIO---KDLNLQHVKLEKTSVEK---IQEAQKESKL 1131
 QY 470 IKDLINHAFOEETGVAVELAEEDDEKIAIKLMLRIEDIKLKGXKONALEFSGFDE 529
 Db 1132 QOKLPKTLSEDKIPATVS-----SDHKLLSSEBKKPEL---LE 1170
 QY 530 RDVPE-----DVAQEMVSGYVCEGDHKTMAKAIK-----DRVSLIK 566
 Db 1171 KSPHPKDKKEQITAEHTTGHTGITEQKVEAPCDKLHEKKQEDVKEDLTGIPQWMSKPE 1230
 QY 567 RKEBQR-----QIVRE---EOEKKKQOESLKOQVQSSASOTGIKOLPAS 610
 Db 1231 KAEBEKTPVPSVSLPRSDHVAAREKTEKEDDSDTSSSQOQSPGLSDTGY-----SS 1285
 QY 611 TGIPPTASTASVTOVE-----PEEP-BADQHQLOQYOQPSISVLSDETVDS---GQG 660
 Db 1286 DGI---SSSIGELPSHIPSDEKULPREPSQKDTISQSPSPSGLATLESTVLISLEQA 1342
 QY 661 SSVETESRVSSQQTV---SYGSQHEQASHTGVTP----- 691
 Db 1343 STLDEKSVKREKLEYETSEQTKQDKHTKPLVTPESYSSDEDELEAIOBGEITIAADSK 1402
 QY 692 GHIPSTVOAQSQPHGVYPP-----SSVAGQSQOQPP-----SSSLTLGVSSSQPIQHP 739
 Db 1403 GGAASSQTDYKEEDGANDTPARQORYDVEDSESSENSPVPRKRAVAGSSSSDEYKRD 1462
 QY 740 QOQOQIOQTAPPOQTVOYSLISOTSTSE-----ATTAPVSOPOAPVULPVYASG--- 789
 Db 1463 SQSGDEEDFIRKQIIEMSADEBASGSHDEFINOLKEISVTESQKKEVKSAXKGTVG 1522
 QY 790 -----KOSTQVGSQVA-----PAEPVAVA----- 808
 Db 1523 KHRMARSSAGYEDADGRHRSHWDDDEFDESPEPKYRETKSQDGEHLAISGGGLRR 1582
 QY 809 -----OPQ----- 811
 Db 1583 FKTIELNSTITSKYSETPEOQKGLIYFDEBELEMEISLTPSPEDRSRGEGSSSLHASFT 1642
 QY 812 -ATOPTLASSVDASHDVASGMSDGENVSSSGRHEG----- 849
 Db 1643 PGTSTSV---SLIDENSDSSPSHKIKLGSQKQKARKHSHGPLPTIEDSBEELREE 1701
 QY 850 -----RTTKRYKRSVSRSHK-----TSRPKRIILVSNKGDVVECOL- 891
 Db 1702 ELKQEKQORELEQOQKSSSKSKKXKODELAQRRRRPKTPPSNLSPIEDASPTIELR 1761
 QY 892 -----ETHNRKMTFKFDLGD-----NPEELATIM-----VNNDFILA 925
 Db 1762 QAAMEEELHRSSCSEYSPSIESDPGCFEISPEKIIIEVQKVYVLPYAVSLYSTDEKILGA 1821
 QY 926 IERSEFVQVREIIEKADML-----SEDSVVEPEGDGL---ESLQK---DDY----- 969

Db 1822 LKEESGQKTLKSAEVEYEMHKTAKSKSFQJASKDEVEFEKESLYGGMILDIYIESLI 1881
 QY 970 ---GFSGS-----QKLEGER-----KOPIPASSM---POQIGIPSSLTQVNHASG 1009
 Db 1882 EDTYNGYVDNLAMQDENSEMEIYIQORGEKKEKIRASBQIYDEFOKL---TDLOEDVYSV- 1936
 QY 1010 RRFIVSPV-ESRLBSKVPESE---ITDVAASTAGSPGMNLHSHASLSLOQAFSL 1064
 Db 1937 -EPLCSIVPOEDIVASSYIIPESHIEIVLDSIVTSTTEKOL-----LDAEAYEEL 1987
 QY 1065 ---RAQMTREGNTAPRPNSSHGPTFPVPPPLSSIAGPPTAAATAPVPATSPENDS 1121
 Db 1988 MKKORWOLT-----PGSSPTQPSDIAFTSDMKVSSIGETADSTLSSTLSAISDS 2040
 QY 1122 TSIVGSEVVP-----TEEGIAG-----VATSTGVVTSGLPIPEVSP- 1159
 Db 2041 -SLSSIALSIDVYKTOHTABEIEDEYLTDAEIQELISHETSLT-----YSSVSG 2094
 QY 1160 -SPVLSVSVSITIPAVVSIITSPSLQVPTSEIYVSSALYPSV----- 1205
 Db 2095 AASILPDSIASLT---SSTSVCITDSSPIDSATVGYVDSAVSKLVBEDIIIAQVPFS 2153
 QY 1206 ----TVSATASAGSSTATPGKPPAVVSOQAAGSTTGALTSTSTTSF---PSTASQI 1259
 Db 2154 TEEYSEVSPVYESVAGATTKPAIASDMDTVHQA---VCLPETAPSFTTVIKPKQYASD 2211
 QY 1260 SIQSSSTSTPTLA-----ETVVVSAHSLDKTSHSSTTGALFASAPSSSSPGAGVS 1312
 Db 2212 TTYVDSIAKDAARKCKGKVERGIIKIHEDSKHLSLDMTRINLTGATISBQPLCVAS 2271
 QY 1313 SVISQPGGLPIV-IPSVIAST-----AAGPTSTPLPOV----- 1351
 Db 2272 VSKYEPASERPAPVTPRVVSKSTSVMSAPALTSKVFSLFRSSSLDSPAQSPSPSP 2331
 QY 1334 -----PLIPO----- 1351
 Db 2332 PPPPPPPPLPPLPPLPKKKSQIQAPMAAPPAVPLVTSVATLESAAVLKHHVVP 2391
 QY 1352 -----PSIPPLVOPVANVA-----VOOTLIHSQOPALLPNOPHTHCPEVDSOTOP 1398
 Db 2392 TKTYTPPPPPVPPKPSIPAGLVFSHRPTEVTVPPIAPKAVAPPLPIAVHKP---AETOP 2448
 QY 1399 KAPQ-----IDIKITLEKL---RSLFSEHS---SGAASHVSLTSIVESTVPIGI 1446
 Db 2449 KPICGLSLSMTINLVSSAEYKIASPTSPSPSHNKSPPRLTKPSQETVYVITLPSBGT 2508
 QY 1447 PTTAV-----APSK-----LLTSTST---CLPPTMLPIGTVALPVTVP 1482
 Db 2509 PTEAITSQAVTSMLEPSPSKQIOPQMPQIFTSMAKAVEIQSMADQSMYI-TGALOITPI 2567
 QY 1483 VTPQAVSTPVSTTSGVAPGTAPASKRPULTAPVLPVGTETLPACTLPSEOLPPEPGLITQ 1542
 Db 2568 TTOSTFEKVPSSKEAVTTEVAKTASVVRPVSGLGVSTIIPREPIYISQPHYRE 2627
 QY 1543 SQQ---PLED-LDAQLRITLSEKXITVTS---VGPVMAA---PTAITAGTPOQGVYS 1592
 Db 2628 NGRHPLIGAVIDL---RLTKVDEIMRDSQMDLSAVSMDARKQWFTSDSG---RPVS 2679
 QY 1593 QYKSGPV-----LATSSGAVFPMGFGFYGVAVADGAQKQKNSSEDAK---S 1636
 Db 2680 TVQPAIINLSTACVADBSLSTVETVAVMTCTATVVSASATDSLVDLGHAMTTPLQLTTS 2739
 QY 1637 VHFESSTSESVLSSSPRESTLVKPEPNGITPIGISDVP-----BSAHTTASEAKSD 1690
 Db 2740 KHFEPAK---VSSQP-----FVPSRDEVPINLSIGTSAHVAWTAATKPV 2782
 QY 1691 TGQPTKV---GRFOVTTAN---KVGRFSVSKEDKITTITKSGPVAAPPFMDQAVLPAV 1746
 Db 2783 TVPPVSVTNGWTLSTSQEPMEIGAVDLSTTKSHRTVVTMD-----ETSGIITTV 2833
 QY 1747 IPKKEP-----ELSEPSHLNGPSSDEPAFLSRVD---D 1779

Db 2834 IEDEKPYDLTLAGRRACVCCDMVYKLPFGSRCTAQP-----PTTLPEDFRGYDDHYQCD 2888
 QY 1780 GSGSPHSFHSK-SLSPSONLSG-----LSNFSNSVYSSNDSEDEDLKLELR 1832
 Db 2889 RSGSYGYGMGMKMSMDTLNSELGLAYKSKSFDVQGTDAVLTSGRVT----- 2943
 QY 1833 LRDHKLKLTODLSQKHEISLYKLGKVPVAVIIPPAALSGRRRPTKSKGSRS 1892
 Db 2944 -----SGEVMYSSK-----TTGPYETRQVIGIGI--STQYSGAR-----MVSSLS 2985
 QY 1893 SSLGKSPQLSGNSGSAASVHQQTLH-----PGNIPSGQNL--LQPLKPSPSD 1946
 Db 2986 SPFGAGSVLRSSN--GVVYSVATPIPTSTFAITTPGSIPTVYADLPLOTIDSVPSLS 3043
 QY 1947 NL-----YSAFTS-----DGAISVPSLSAPGCGTSTTVGATVN 1981
 Db 3044 TLQGNQPLPFSYSLTTLTAARKDASTIIDITETGPLETLISIAIPIPLIPATAS 3098

RESULT 14

MAK4 HUMAN STANDARD; FRT; 1239 AA.

AC 095819; 075172; Q9NST7;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Mtoogen-activated protein kinase kinase kinase 4 (EC 2.7.1.37)
 DE (MAPK/ERK kinase kinase 4) (MEK kinase 4) (MEKK 4)
 DE (HRK/GCK-like kinase HGK) (Nck interacting kinase).
 GN MAP4K4 OR HGK OR NIK OR KTA0687.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid=9606;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 4 AND 5), FUNCTION, AND TISSUE
 RP SPECIFICITY.
 RP TISSUE=Macrophage;
 RX MEDLINE=99107863; PubMed=9890973;
 RT "A novel human STE20-related protein kinase, HGK, that specifically
 RT activates the c-Jun N-terminal kinase signaling pathway.";
 RT J. Biol. Chem. 274.2118-2125(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RP TISSUE=Glioblastoma;
 RX MEDLINE=22499944; PubMed=12612079;
 RT Wright J.H., Wang X., Manning G., Lamere B.J., Le P., Zhu S.,
 RA Khatri D., Flanagan P.M., Buckley S.D., Whyte D.B., Howlett A.R.,
 RA Bischoff J.R., Lipsen K.E., Jallat B.;
 RT "The STE20 kinase HGK is broadly expressed in human tumor cells and
 RT can modulate cellular transformation, invasion, and adhesion.";
 RT Mol. Cell. Biol. 23:2068-2082(2003).
 RN [3]
 RP SEQUENCE OF 38-1239 FROM N.A. (ISOFORM 2).
 RP TISSUE=Brain;
 RX MEDLINE=98403880; PubMed=9734811;
 RT Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
 RA Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. X.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RT DNA Res. 5:169-176(1998).
 RN [4]
 RP SEQUENCE OF 262-1239 FROM N.A. (ISOFORM 3).
 RP TISSUE=Testis;
 RX Duesterhoeft A., lauber J., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 700-1239 FROM N.A.
 RP TISSUE=Brain;
 RX Saito T., Seki N., Hori T.;
 RT "Isolation, expression profile and chromosome assignment of a novel

RT serine/threonine kinase gene.";
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP INTERACTION WITH SPG3A.
 RX MEDLINE=22275851; PubMed=1238798;
 RA Juan Z., Zhang Y., Liu A., Man Y., Cheng L., Hu G.;
 RT "A novel GTP-binding protein HGBP3 interacts with NIK/HGK.";
 RL FEBS Lett. 530:233-238(2002).
 CC -1- FUNCTION: Serine/threonine kinase that may play a role in the
 CC response to environmental stress and cytokines such as TNF-alpha.
 CC Appears to act upstream of the c-Jun N-terminal pathway.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- COFACTOR: Magnesium.
 CC -1- SUBUNIT: Interacts with the SH3 domain of the adapter proteins Nck
 CC (by similarity). Binds, via its CNH regulatory domain, to the N-
 CC terminal region of SPG3A.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=5;
 CC Name=1; Synonyms=Tumor-associated;
 CC IsoId=O95819-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O95819-2; Sequence=VSP_007054, VSP_007057;
 CC Name=3;
 CC IsoId=O95819-3; Sequence=VSP_007056, VSP_007057;
 CC Name=4; Synonyms=HGK-S;
 CC IsoId=O95819-4; Sequence=VSP_007054, VSP_007055, VSP_007057,
 CC VSP_007058;
 CC Name=5; Synonyms=HGK-L;
 CC IsoId=O95819-5; Sequence=VSP_007054, VSP_007055, VSP_007056,
 CC VSP_007057;
 CC -1- TISSUE SPECIFICITY: Appears to be ubiquitous, expressed in all
 CC tissue types examined. Isoform 5 appears to be more abundant in
 CC the brain, isoform 4 is predominant in the liver, skeletal muscle
 CC and placenta.
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC STE20 subfamily.
 CC -1- SIMILARITY: Contains 1 CNH domain.
 CC -1- CAUTION: Ref.4 sequence differs from that shown due to a
 CC frameshift in position 1151.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF096300; AAD16137.1; -;
 CC EMBL; AY212247; AAO32626.1; -;
 CC EMBL; AB014587; BAA31662.1; -;
 CC EMBL; AL137755; CAB70907.1; ALT_FRAME.
 CC EMBL; AB013385; BAA33714.1; -;
 CC Genew; HGNC:6866; MAP4K4.
 CC MIM; 604665; -;
 CC GO; GO:0005524; F-ATP binding; IDA.
 CC GO; GO:0004674; F-protein serine/threonine kinase activity; IDA.
 CC GO; GO:0006468; P-protein amino acid phosphorylation; IDA.
 CC GO; GO:0007243; P-protein kinase cascade; IDA.
 CC GO; GO:0006950; P:response to stress; IDA.
 CC InterPro; IPR001180; Cytrom.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR008271; Ser_thr_kin_AS.
 CC InterPro; IPR002290; Ser_thr_kinase.
 CC InterPro; IPR001245; Tyr_pkinase.
 CC Pfam; PF00780; CNH; 1.
 CC Pfam; PF00069; Pkinase; 1.
 CC ProDom; PD000601; Prot_kinase; 1.
 CC SMART; SM00036; CNH; 1.
 CC SMART; SM00220; STK; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.

Blank sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 00:59:43 ; Search time 195 Seconds
(without alignments)
3456.138 Million cell updates/sec

Title: US-10-010-720-14
Sequence: 1 MEGGAALKKSTPGSLFLSP.....NISNLQKSIKNPPGSLNLTFT 2136
Perfect score: 10812

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP TREMBL 25: *
2: sp archaea: *
3: sp bacteria: *
4: sp fungi: *
5: sp human: *
6: sp invertebrate: *
7: sp mammal: *
8: sp organelle: *
9: sp phage: *
10: sp plant: *
11: sp rodent: *
12: sp virus: *
13: sp vertebrate: *
14: sp unclassified: *
15: sp virus: *
16: sp bacteriophage: *
17: sp archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10677	96.8	2382	4	Q9H4A3
2	10435	96.5	2066	4	O15052
3	9204	85.1	2126	11	Q9J1H7
4	3408	31.5	670	4	Q9P1S9
5	2630	24.3	2296	4	O9Y3S1
6	2615	22.8	2219	4	O9C0A3
7	2463.5	22.4	1800	4	O8TCX6
8	2439	22.6	1743	4	O9BYP7
9	2126	19.7	1243	4	O96J92
10	2083	19.3	1222	11	O810H4
11	2082	19.3	1210	11	O811R5
12	2068	19.1	1222	11	O80UR6
13	1974	18.3	1393	4	O8N673
14	1911.5	17.7	1048	11	O80XN2
15	1890.5	17.5	1552	11	O7TPK6
16	1804.5	16.7	779	4	Q9H3P4

17	1682.5	15.6	417	11	O8K2Q3	O8K2Q3 mus musculus
18	1655.5	15.3	1839	5	O18657	O18657 caenorhabditis
19	1610	14.9	1678	5	O81127	O81127 caenorhabditis
20	1507.5	13.9	954	11	O80XB5	O80XB5 mus musculus
21	1426	13.2	1557	11	O811F2	O811F2 mus musculus
22	1362.5	12.6	663	4	O8N8X3	O8N8X3 homo sapiens
23	1238	11.5	2308	5	O9VP17	O9VP17 drosophila
24	1030.5	9.5	1213	4	O9HCK6	O9HCK6 homo sapiens
25	945	8.7	212	4	O9BY55	O9BY55 homo sapiens
26	915	8.5	173	4	O96CZ6	O96CZ6 homo sapiens
27	856	7.9	176	11	O7TPR2	O7TPR2 mus musculus
28	768	7.1	591	10	O84R51	O84R51 medicago sa
29	761.5	7.0	563	10	O944Q0	O944Q0 arabidopsis
30	761	7.0	608	10	O9FHY4	O9FHY4 arabidopsis
31	759.5	7.0	550	10	O8S8Y7	O8S8Y7 arabidopsis
32	755.5	7.0	557	10	O8LS12	O8LS12 arabidopsis
33	755.5	7.0	666	10	O9PDV6	O9PDV6 fagus sylvia
34	752.5	7.0	567	10	O8S8Y8	O8S8Y8 arabidopsis
35	750.5	6.9	578	10	O9LUV8	O9LUV8 arabidopsis
36	746	6.9	524	10	O8RXE5	O8RXE5 arabidopsis
37	745.5	6.9	516	10	O9STK6	O9STK6 arabidopsis
38	742.5	6.9	568	10	O8S8Y9	O8S8Y9 arabidopsis
39	740.5	6.8	571	10	O9LSA2	O9LSA2 arabidopsis
40	715.5	6.6	539	10	O9CSN1	O9CSN1 arabidopsis
41	709.5	6.6	542	10	O9M9B6	O9M9B6 arabidopsis
42	702.5	6.5	542	10	O94HK9	O94HK9 oryza sativa
43	702.5	6.5	542	10	O7XB94	O7XB94 oryza sativa
44	702.5	6.5	547	10	O9SCU5	O9SCU5 arabidopsis
45	702.5	6.5	549	10	O9XF55	O9XF55 arabidopsis

ALIGNMENTS

RESULT 1

Q9H4A3 PRELIMINARY; PRT; 2382 AA.
ID Q9H4A3
AC Q9H4A3; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN WNK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Verissimo F., Jordan P.;
RT "A novel human protein kinase gene, WNK1: cDNA cloning, expression and genomic structure."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ296290; CAC15059.1; -
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD00001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP_binding; Transferase;
KW Serine/threonine-protein kinase; ATP_binding; Transferase.
SQ SEQUENCE 2382 AA; 250755 MW; ADDA6C86554B7CA CRC64;

Query Match 98.8%; Score 10677; DB 4; Length 2382;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;

QY 1 MSGAAEKOSSTPGSLFLSPAPAPKXSSSDSVGEKLGAAAADAVTGRTEYRRRRRT 60
 Db 1 MSGAAEKOSSTPGSLFLSPAPAPKXSSSDSVGEKLGAAAADAVTGRTEYRRRRRT 60
 QY 61 MKDORGAATTTTTEHRRFRFRSVICDSNAKLELPLSLPSPISPAVQSAPEBPH 120
 Db 61 MKDORGAATTTTTEHRRFRFRSVICDSNAKLELPLSLPSPISPAVQSAPEBPH 120
 QY 121 REETATATSOVAOOPPAAPAGEOAVVGPAPSTVPSTSKDRVSOQSVGSKEEPP 180
 Db 121 REETATATSOVAOOPPAAPAGEOAVVGPAPSTVPSTSKDRVSOQSVGSKEEPP 180
 QY 181 ARSGGGGSAKXPOEERSQOQDDIELEETKAVGMSNDGRFLKFDIEIGRSEFKTYKGLD 240
 Db 181 ARSGGGGSAKXPOEERSQOQDDIELEETKAVGMSNDGRFLKFDIEIGRSEFKTYKGLD 240
 QY 241 TETTVAVMCELQORUKLTCSRORKEBAEMLKGIQHPNIVRFYDSMESTYKXKCIYLV 300
 Db 241 TETTVAVMCELQORUKLTCSRORKEBAEMLKGIQHPNIVRFYDSMESTYKXKCIYLV 300
 QY 301 TELMTSGTLTKYLRKFYMKIKVLRSWCROILKGIQFHTPTPTPIHDKIKDNIPTGP 360
 Db 301 TELMTSGTLTKYLRKFYMKIKVLRSWCROILKGIQFHTPTPTPIHDKIKDNIPTGP 360
 QY 361 TGSYKIGDLGLATLKRAFAKSVIGTPEFMAPEMYEERKYDESVDVYAFGMCMLEMAISEY 420
 Db 361 TGSYKIGDLGLATLKRAFAKSVIGTPEFMAPEMYEERKYDESVDVYAFGMCMLEMAISEY 420
 QY 421 PYSCQNAAOIYRVTSGVKPASFDAKVAIPVXEKIEBCTIKQNDERYSIKDLINHAFFQ 480
 Db 421 PYSCQNAAOIYRVTSGVKPASFDAKVAIPVXEKIEBCTIKQNDERYSIKDLINHAFFQ 480
 QY 481 BETGVARELAEDDEGEKIAIKLMLRIEDIKLKJCKYKONBEATERSPLERVPEDVAQEM 540
 Db 481 BETGVARELAEDDEGEKIAIKLMLRIEDIKLKJCKYKONBEATERSPLERVPEDVAQEM 540
 QY 541 VESGYVEBGHKTMAKAIKORVSLIKRREGORQVREOEKKOEBSLKQOVQOSASQ 600
 Db 541 VESGYVEBGHKTMAKAIKORVSLIKRREGORQVREOEKKOEBSLKQOVQOSASQ 600
 QY 601 TGIKOLPSASTGIPTAATTSASVSTQVEPEPEADHQOQLOYOQPSISVSDGTVDGSGG 660
 Db 601 TGIKOLPSASTGIPTAATTSASVSTQVEPEPEADHQOQLOYOQPSISVSDGTVDGSGG 660
 QY 661 SSVFTEBRSVSOQTVSYSQHEQAHSTGYTPGHPSTVVOAOSQPHGVYPPSSVAQGGSGG 720
 Db 661 SSVFTEBRSVSOQTVSYSQHEQAHSTGYTPGHPSTVVOAOSQPHGVYPPSSVAQGGSGG 720
 QY 721 QBSSSSLTGVSSOPIOHPOOQOGIOQTAPPOQTVOYSLQSTSTSEATTAQEVSOQAP 780
 Db 721 QBSSSSLTGVSSOPIOHPOOQOGIOQTAPPOQTVOYSLQSTSTSEATTAQEVSOQAP 780
 QY 781 QVLPQVSAGKQ----- 791
 Db 781 QVLPQVSAGKQ----- 791
 QY 792 ----- 791
 Db 792 ----- 791
 QY 841 VSGOPISTPHVSTAQTFSSLPITMAAGITOPLLTLASSATTAIPGVSTVPSQLPTLL 900
 Db 841 VSGOPISTPHVSTAQTFSSLPITMAAGITOPLLTLASSATTAIPGVSTVPSQLPTLL 900
 QY 901 QPVTOQPSQVHPOLLOPAVOSMGIPANLGAAEVPPLSGDVLVYQGFPRPLPQYPGDSNI 960
 Db 901 QPVTOQPSQVHPOLLOPAVOSMGIPANLGAAEVPPLSGDVLVYQGFPRPLPQYPGDSNI 960
 QY 961 ABSSNVAVCIHSTVLSPPMTEVLATPGYFPTVQPYVESNLVPMGVGQGVQVYQCPG 1020
 Db 961 ABSSNVAVCIHSTVLSPPMTEVLATPGYFPTVQPYVESNLVPMGVGQGVQVYQCPG 1020
 QY 792 -----STQVSOVAPAPBPVAVAOQATOPTTLASVSDVASHSDVASGMS 834
 Db 792 -----STQVSOVAPAPBPVAVAOQATOPTTLASVSDVASHSDVASGMS 834
 QY 1021 GSLAQAPTTSSQOAVLESTQGVSVAPAPVAVAOQATOPTTLASVSDVASHSDVASGMS 1080
 Db 1021 GSLAQAPTTSSQOAVLESTQGVSVAPAPVAVAOQATOPTTLASVSDVASHSDVASGMS 1080

QY 835 DGENVVPSSGGRHGGRTTKRHVRKSVRSRREKTSRPEKRLILVNSKNGRVVECOLETH 894
 Db 1061 DGENVVPSSGGRHGGRTTKRHVRKSVRSRREKTSRPEKRLILVNSKNGRVVECOLETH 1140
 QY 895 NRKWTFFKFDLDGNDPEELATIMVNDPLATERSSFPVDQVEIIEKADEMISEVSEPP 954
 Db 1141 NRKWTFFKFDLDGNDPEELATIMVNDPLATERSSFPVDQVEIIEKADEMISEVSEPP 1200
 QY 955 EGDQGLBSLQCKDQVDFSGSQKLEGEFKOPIPASSMPOOIGIPTSSLQOVHSAGRFV 1014
 Db 1201 EGDQGLBSLQCKDQVDFSGSQKLEGEFKOPIPASSMPOOIGIPTSSLQOVHSAGRFV 1260
 QY 1015 SPVESRRLRESKVPSELTIDVVAASTAOSPQNLSSHASLSLQAFSELRAQMTBGN 1074
 Db 1261 SPVESRRLRESKVPSELTIDVVAASTAOSPQNLSSHASLSLQAFSELRAQMTBGN 1320
 QY 1075 TAPNFSHTGTFPVPVPEFLSSIGVPTTAATKVPVPTSSPPNDISTSVIQSEVTYTE 1134
 Db 1321 TAPNFSHTGTFPVPVPEFLSSIGVPTTAATKVPVPTSSPPNDISTSVIQSEVTYTE 1380
 QY 1135 EGIAGVASTGVTSGGLPIPPVSESPVSSVSSITIPAVVSISSTPSLQVPTSEI 1194
 Db 1381 EGIAGVASTGVTSGGLPIPPVSESPVSSVSSITIPAVVSISSTPSLQVPTSEI 1440
 QY 1195 VVSSTALYPSVTGATSAAGGSTATPGKPPAVVSOQAAGSTTVGATLISVSTTSPS 1254
 Db 1441 VVSSTALYPSVTGATSAAGGSTATPGKPPAVVSOQAAGSTTVGATLISVSTTSPS 1500
 QY 1255 TASQSLTOLSSSTPTLAETVVVVAHSLDKTSHSTTGAFSLASASSSSPAGVSSY 1314
 Db 1501 TASQSLTOLSSSTPTLAETVVVVAHSLDKTSHSTTGAFSLASASSSSPAGVSSY 1560
 QY 1315 ISOPGGLHPIVIBSVIASTPILPOAAGTSTPILLQVPSIPPLVQPVANYPVQOULIHS 1374
 Db 1561 ISOPGGLHPIVIBSVIASTPILPOAAGTSTPILLQVPSIPPLVQPVANYPVQOULIHS 1620
 QY 1375 QOPALLPNOPHHCPEVSDTOPKAPGIDDKYLEKKLASLSEHSSSGAQAHSVLET 1434
 Db 1621 QOPALLPNOPHHCPEVSDTOPKAPGIDDKYLEKKLASLSEHSSSGAQAHSVLET 1680
 QY 1435 SLVIESVTPEGIPPTVAAPSKLSTSTGCLPPTNLPLGVVALPVPVVPVQGVSTVST 1494
 Db 1681 SLVIESVTPEGIPPTVAAPSKLSTSTGCLPPTNLPLGVVALPVPVVPVQGVSTVST 1740
 QY 1495 TTSQVKGTAESPRLTKAPVLPVGTLEPAQTLPSQPLPFPSPSILQSQPLEDDAQL 1554
 Db 1741 TTSQVKGTAESPRLTKAPVLPVGTLEPAQTLPSQPLPFPSPSILQSQPLEDDAQL 1800
 QY 1555 RRTLSPEXITTVSANGPVMAAPTAATTEAGTOPQKGVSOYKEGPVLTSSGAGVFKXGRF 1614
 Db 1801 RRTLSPEXITTVSANGPVMAAPTAATTEAGTOPQKGVSOYKEGPVLTSSGAGVFKXGRF 1860
 QY 1615 QVSVADGAQKQKXSEDAKSVHFESSTSESSVSSPESLTVPBEPGITTPIGSSD 1674
 Db 1861 QVSVADGAQKQKXSEDAKSVHFESSTSESSVSSPESLTVPBEPGITTPIGSSD 1920
 QY 1675 VPBSAKHTTASBAKSDTGPQTVKGRFOVTTANKVGRFVSXKTEDKITDTKKEGPVAPSP 1734
 Db 1921 VPBSAKHTTASBAKSDTGPQTVKGRFOVTTANKVGRFVSXKTEDKITDTKKEGPVAPSP 1980
 QY 1735 FMDLEQAVLPVAVIPKKEKPELSEPSHLNGSSDPKAPFLSRVYDSDGSGPHSHQLSKSG 1794
 Db 1981 FMDLEQAVLPVAVIPKKEKPELSEPSHLNGSSDPKAPFLSRVYDSDGSGPHSHQLSKSG 2040
 QY 1795 LPSQMLSGSLNSFNSSVMSDNESDIEDDKLELRLRDLKHEIDLOJROKHETES 1854
 Db 2041 LPSQMLSGSLNSFNSSVMSDNESDIEDDKLELRLRDLKHEIDLOJROKHETES 2100
 QY 1855 LYTXYKGPVPVAVITPPAAPLSGRRRRPTKSKSKSSRSSSLGNKSPOLSGNLSGQSAASV 1914
 Db 2101 LYTXYKGPVPVAVITPPAAPLSGRRRRPTKSKSKSSRSSSLGNKSPOLSGNLSGQSAASV 2160
 QY 1915 LHPQOTLHPPGNIPESQGNQLQPLKPSBSSDNLXSAFTSDGALSVPSLSAPGGTSSSTN 1974


```

DB      2161 LHPQOTLHPGNIIPSSGQWLLQPLKPPSSSNLVSAFSDGASIVPSLSADGQTSSTN 2220
QY      1975 TVGATVNSQAQAQAPPATSSRRKGTFTDHLKLVNMAADANLSSRRSKHNNYEGG 2034
DB      2221 TVGATVNSQAQAQAPPATSSRRKGTFTDHLKLVNMAADANLSSRRSKHNNYEGG 2280
QY      2035 MARKESAPQOLCISMTSNLGGSAFISASATSIGHFTKSMCPPOYGFPATPEGAQMSG 2094
DB      2281 MARKESAPQOLCISMTSNLGGSAFISASATSIGHFTKSMCPPOYGFPATPEGAQMSG 2340
QY      2095 GGPAPQPLGQFQVGTASLQNFNISNLQKSNISNPPGSNLRTT 2136
DB      2341 GGPAPQPLGQFQVGTASLQNFNISNLQKSNISNPPGSNLRTT 2382

RESULT 2
015052 PRELIMINARY; PRT; 2066 AA.
ID      015052;
AC      015052;
DT      01-JUN-1998 (TrEMBLrel. 05, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein KIAA0344 (Fragment).
GN      KIAA0344.
OS      Homo sapiens (human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SOURCE FROM N.A.
RC      TISSUE=Brain;
RX      MEDLINE=97349984; PubMed=9205841;
RA      Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
RA      Tanaka A., Kotani H., Nomura N., Ohara O.;
RT      "Prediction of the coding sequences of unidentified human genes. VII.
RT      The complete sequences of 100 new cDNA clones from brain which can
RT      code for large proteins in vitro.";
RL      DNA Res. 4:141-150(1997).
RL      EMBL; AB002342; BAA20802.2; -.
DR      Genew; HGNC:14540; PRKNKL.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR      GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR      GO; GO:0004678; F:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR000719; Prot_kinase.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      InterPro; IPR008271; Ser_thr_kinase.
DR      InterPro; IPR001245; Tyr_kinase.
DR      Pfam; PF00069; Pkinase; 1.
DR      ProDom; PD000001; Prot_kinase; 1.
DR      SMART; SM00220; S_TKC; 1.
DR      SMART; SM00219; TyTK; 1.
DR      PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR      PROSITE; PS00108; PROTEIN KINASE ST; 1.
KM      Hypothetical protein; ATP-binding; Transferase.
FT      NON_TER
SQ      SEQUENCE 2066 AA; 218551 MW; EC59F36363D45D3 CR664;

Query March 96.5%; Score 10435; DB 4; Length 2066;
Best local similarity 99.9%; Pred. No. 0;
Matches 2065; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```

```

DB      121 SAKPEBRSQQDDIELEETKAVGMSNDGRFLKDIIEIGRSFKTVYKGLDTETTVEVA 180
QY      249 WCELDRLKLTSEORRFKEEAEMLKGLQHPNIVREFYDSWESIVKSKCIVLVEMLTSGT 308
DB      181 WCELDRLKLTSEORRFKEEAEMLKGLQHPNIVREFYDSWESIVKSKCIVLVEMLTSGT 240
QY      309 LKTVLKRPMKIKVLRSMCRQILKGLQFLHTRTPPIIHRDKCNPIFTGTGSKYKID 368
DB      241 LKTVLKRPMKIKVLRSMCRQILKGLQFLHTRTPPIIHRDKCNPIFTGTGSKYKID 300
QY      369 LGLATLKRASTFKSVYIGTFEPFAPMEYBEKKDESVDVAFGCMLEMATSEYSECCQA 428
DB      301 LGLATLKRASTFKSVYIGTFEPFAPMEYBEKKDESVDVAFGCMLEMATSEYSECCQA 360
QY      429 AQTIVRVTSVGPASPDKVAIPVEVIEIGCIRONKDERYSIKDLINFAFPOEETGVRE 488
DB      361 AQTIVRVTSVGPASPDKVAIPVEVIEIGCIRONKDERYSIKDLINFAFPOEETGVRE 420
QY      489 LAEEDGEEKIATKMLRIBEDIKKLGKYNDAIEFSFLEEDVPEVDAQEAVESGYCE 548
DB      421 LAEEDGEEKIATKMLRIBEDIKKLGKYNDAIEFSFLEEDVPEVDAQEAVESGYCE 480
QY      549 GGHKTMAKAIKORVSLIKRREORQVREOEKKQEESLQQVQSSASQIGIKQLPS 608
DB      481 GGHKTMAKAIKORVSLIKRREORQVREOEKKQEESLQQVQSSASQIGIKQLPS 540
QY      609 ASNGIPTATTSASVSTOVEPEPEADQHQLOQYQOPSTISVLSGDTVDSGQSSVFTESR 668
DB      541 ASNGIPTATTSASVSTOVEPEPEADQHQLOQYQOPSTISVLSGDTVDSGQSSVFTESR 600
QY      669 VSSQQTVSYGSGHQEQAHSSTGVPGHISTVQAQSGPHGYVPPSSVAQSGSQGPPSSSLT 728
DB      601 VSSQQTVSYGSGHQEQAHSSTGVPGHISTVQAQSGPHGYVPPSSVAQSGSQGPPSSSLT 660
QY      729 GUSSSQPIQHPOQOQGIQCTAPPOQVQVYSLQSTSTSSKTTAKOPVOPAPQVLRQVSA 788
DB      661 GUSSSQPIQHPOQOQGIQCTAPPOQVQVYSLQSTSTSSKTTAKOPVOPAPQVLRQVSA 719
QY      789 GKSTQGVSGVAPAEPAVAVQAPQATOPTTLASVDSASHSVASGMSGNENVSSSGRHE 848
DB      720 GKSTQGVSGVAPAEPAVAVQAPQATOPTTLASVDSASHSVASGMSGNENVSSSGRHE 779
QY      849 GRTTKHRYKRSVRSRHEKTSRPKLRILNVSKGDRVVECOLETHNRKAVTFKFDLDC 908
DB      780 GRTTKHRYKRSVRSRHEKTSRPKLRILNVSKGDRVVECOLETHNRKAVTFKFDLDC 839
QY      909 NPEBIATIMVNDPILAIERESFVDQVREIIEKADEMLSRDVSVEPRDQGLESLQKDP 968
DB      840 NPEBIATIMVNDPILAIERESFVDQVREIIEKADEMLSRDVSVEPRDQGLESLQKDP 899
QY      969 YGFSGQKLEGEFPKQPIPASSMPQOIGIPSSILQVVAHAGRRITVSPVESRLRESKV 1028
DB      900 YGFSGQKLEGEFPKQPIPASSMPQOIGIPSSILQVVAHAGRRITVSPVESRLRESKV 959
QY      1029 PSEITDTVAASTAGSPGMNLSHSAISLQQAPELRAQMTBEPNAPNPFSTGTTFP 1088
DB      960 PSEITDTVAASTAGSPGMNLSHSAISLQQAPELRAQMTBEPNAPNPFSTGTTFP 1019
QY      1089 VVPEFLSSINAGVPTTAATAVPVPAATSSPPNDISTSVQSEVTVTEEGIAGVASTGVVT 1148
DB      1020 VVPEFLSSINAGVPTTAATAVPVPAATSSPPNDISTSVQSEVTVTEEGIAGVASTGVVT 1079
QY      1149 SGGPIPEVSESPVSSVSVSSITIPAVVSISSITSPSLQVPTSEIIVSSSTALPVSTVS 1208
DB      1080 SGGPIPEVSESPVSSVSVSSITIPAVVSISSITSPSLQVPTSEIIVSSSTALPVSTVS 1139
QY      1209 ATSAAGGSTATPEPKPPAVVSGQAAGSTYGAULTSVSTTSPSTASQSLQSSSTS 1268
DB      1140 ATSAAGGSTATPEPKPPAVVSGQAAGSTYGAULTSVSTTSPSTASQSLQSSSTS 1199
QY      1269 TPTLAETVTVSAHSLDTSHTSTTGIAFLSLAPSSSSPPGAVSGYIQPGHILVLPVS 1328

```

Db 1200 TPTLAETVAVVSAHSDKTSHTSSTGLAFLSAPSSSSSPGAGVSSYISQPGGLHPLVIPS 1259
 QY 1329 VIATPILIPQAGPTSTPLPQVPSIPVLVQVAVVAVVAVQOTLHISOPPALPMPQPTH 1388
 Db 1260 VIATPILIPQAGPTSTPLPQVPSIPVLVQVAVVAVVAVQOTLHISOPPALPMPQPTH 1319
 QY 1389 CREVSDTQPKAPGIDDKITLSEKRLSLFSEHSSSGAQAHSVLETSVIESTVPGIPT 1448
 Db 1320 CPEVSDTQPKAPGIDDKITLSEKRLSLFSEHSSSGAQAHSVLETSVIESTVPGIPT 1379
 QY 1449 TAVASKLLTSTSTSCLPPTNLPLGTVALPLVTVVTPQVSTPSTTSVGVKPGTAPSKP 1508
 Db 1380 TAVASKLLTSTSTSCLPPTNLPLGTVALPLVTVVTPQVSTPSTTSVGVKPGTAPSKP 1439
 QY 1509 PLTAAPVLPGVTELPAGTLPSBOLPFPFGPSITQSQCPLELDLADLRRTLSPEXITVISA 1566
 Db 1440 PLTAAPVLPGVTELPAGTLPSBOLPFPFGPSITQSQCPLELDLADLRRTLSPEXITVISA 1499
 QY 1569 VGPVMAAPTAITENGTPQKGVGVKGPVLATSSGAVFRMGKRFQVSVADGAQKXGK 1628
 Db 1500 VGPVMAAPTAITENGTPQKGVGVKGPVLATSSGAVFRMGKRFQVSVADGAQKXGK 1558
 QY 1629 NKSEDAKSVFESSSTSESVSSSPSTLYKPEPNCITIGISDVPEAHKTTASPAK 1668
 Db 1559 NKSEDAKSVFESSSTSESVSSSPSTLYKPEPNCITIGISDVPEAHKTTASPAK 1618
 QY 1689 SDTQPKTVKGFQVYTTANKVGRFSVSKTEDKIDTKKEGVAAPPPMDLEQAVLPAVTP 1748
 Db 1619 SDTQPKTVKGFQVYTTANKVGRFSVSKTEDKIDTKKEGVAAPPPMDLEQAVLPAVTP 1678
 QY 1749 KKEPELSEBHLNPGSSDPEAFLSRVDVDSGSPHSPHQLSKSLPSQNLSSQSLNSF 1808
 Db 1679 KKEPELSEBHLNPGSSDPEAFLSRVDVDSGSPHSPHQLSKSLPSQNLSSQSLNSF 1738
 QY 1809 NSSYMSDNESEDIEDDLKELRLRDKHLKEIODLSROKHEIESUYTLKGVKPPAVIT 1868
 Db 1739 NSSYMSDNESEDIEDDLKELRLRDKHLKEIODLSROKHEIESUYTLKGVKPPAVIT 1798
 QY 1869 PPAAPLGGRRRRPTKSGKSSRSSSLGANKSPOLSGMLSGQSAASVLAHPQOTLHPNIP 1928
 Db 1799 PPAAPLGGRRRRPTKSGKSSRSSSLGANKSPOLSGMLSGQSAASVLAHPQOTLHPNIP 1858
 QY 1929 ESGONOLQPLKSPSSDNLISAFSTSGAISVPSLSAPGQSTSTVTVGATVNSQAQAQ 1988
 Db 1859 ESGONOLQPLKSPSSDNLISAFSTSGAISVPSLSAPGQSTSTVTVGATVNSQAQAQ 1918
 QY 1989 PPAATSSRKGTFTDDLHKLVNNAARDAMNLGRRGSGHNNYEGPGMARKEFAPGULCIS 2048
 Db 1919 PPAATSSRKGTFTDDLHKLVNNAARDAMNLGRRGSGHNNYEGPGMARKEFAPGULCIS 1978
 QY 2049 MTSNLGSAPIISAASATSLGHFTKSCPPQOYGPATPFGAOWSGTGPAQPLGQFOFV 2108
 Db 1979 MTSNLGSAPIISAASATSLGHFTKSCPPQOYGPATPFGAOWSGTGPAQPLGQFOFV 2038
 QY 2109 GTSALQNFNINSLQKSIINPDSNLT 2136
 Db 2039 GTSALQNFNINSLQKSIINPDSNLT 2066

RP SEQUENCE FROM N.A.
 RX MEDLINE=20287550; PubMed=10828064;
 RA Xu B., English J.M., Wilsbacher J.L., Stippec S., Goldsmith E.J.,
 RT "WNKI, a Novel Mammalian Serine/Threonine Protein Kinase Lacking the
 RT Catalytic Lysine in Subdomain II".
 RL J. Biol. Chem. 275:16795-16801(2000).
 CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF227741; AAF74258.1; -
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot. Kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot. Kinase; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 2126 AA; 225216 MW; 7137BC45A54402F CRC64;
 Query Match 85.1%; Score 9204; DB 11; Length 2126;
 Best Local Similarity 85.6%; Pred. No. 0;
 Matches 1838; Conservative 83; Mismatches 193; Indels 34; Gaps 15;
 QY 1 MSGGAAGKSGSTGSLFLSPAPAPKNGSSSDSSVGBKLGAAADAVTGRTEYRRRRHT 60
 Db 1 MSDGTAEKSGGTG--FLSPAPAPKNGSSSDSSVGBKLGAAADAVTGRTEYRRRRHT 58
 QY 61 MDKDSRGAATTTTTHRRFRFRSVICDSNATALEPLSLPLSLOPSPAPAVPOGAPPEP 120
 Db 59 MDKDSRGAATTTTTHRRFRFRSVICDSNATALEPLSLPLSLOPSPAPAVPOGAPPEP 118
 QY 121 REETVATATSOVAOQPPAAAPGEQAVAPASTVSSSTSKRPVSGPSVSGKEPPP 180
 Db 119 REETVATATSOVAOQPPAAAPGEQAVAPASTVSSSTSKRPVSGPSVSGKEPPP 178
 QY 181 AR--SSGGGSAKEPQERSQDDIELETKAVGMSNDRFLKFDIEIRGSEKTYVG 238
 Db 179 SRGSGSGGASAEPEERQDDIELETKAVGMSNDRFLKFDIEIRGSEKTYVG 236
 QY 239 LDLETTVEVAMCELOBRKLTYSERORPKFEAEMLKGLQHNIVYRFPDSWSTYKGGKCI 298
 Db 239 LDLETTVEVAMCELOBRKLTYSERORPKFEAEMLKGLQHNIVYRFPDSWSTYKGGKCI 298
 QY 239 LDLETTVEVAMCELOBRKLTYSERORPKFEAEMLKGLQHNIVYRFPDSWSTYKGGKCI 298
 Db 239 LDLETTVEVAMCELOBRKLTYSERORPKFEAEMLKGLQHNIVYRFPDSWSTYKGGKCI 298
 QY 299 LVTELMTSGTLKTYLKRFPVKIKIVLRWCROTLKGLQFLHTRTPILIHDLKCDNIFIT 358
 Db 299 LVTELMTSGTLKTYLKRFPVKIKIVLRWCROTLKGLQFLHTRTPILIHDLKCDNIFIT 358
 QY 359 GPTGSVKIDGLGATLKRAKSPAKSVITGTEPNAPPEMYEKEYDESVDVYAFGCMLEMAT 418
 Db 359 GPTGSVKIDGLGATLKRAKSPAKSVITGTEPNAPPEMYEKEYDESVDVYAFGCMLEMAT 418
 QY 419 EYPSQCNAAQIYRRTVSGVAPASFDKVAIPVKKIIEGCTRONDERYSITDLNNAF 478
 Db 419 EYPSQCNAAQIYRRTVSGVAPASFDKVAIPVKKIIEGCTRONDERYSITDLNNAF 478
 QY 479 FQETGVARELAEDDEGKIAIKLMLRIEDIKLKXKYKDNEAIEBSPLEDERVPEVDAQ 538
 Db 479 FQETGVARELAEDDEGKIAIKLMLRIEDIKLKXKYKDNEAIEBSPLEDERVPEVDAQ 538
 QY 539 EWEVSGYVCEGDKHTMAKAIKRVSLIRKKRQROLVREOEKKQOESSLSKQVE--QSS 597
 Db 539 EWEVSGYVCEGDKHTMAKAIKRVSLIRKKRQROLVREOEKKQOESSLSKQVE--QSS 598
 QY 598 ASQTGIKOLPASASTGIPTASTSASVSTQVEBEPEADHQOLQVQPSISVLSGTVDS 657
 Db 598 ASQTGIKOLPASASTGIPTASTSASVSTQVEBEPEADHQOLQVQPSISVLSGTVDS 658
 QY 658 GGGSSVFTESTRVSSQQTYSYSGOHEOASHGTVPGHIPSTVQVAGSPGAVYPPSSVAQO 717
 Db 658 GGGSSVFTESTRVSSQQTYSYSGOHEOASHGTVPGHIPSTVQVAGSPGAVYPPSSVAQO 718

QY 718 SOGQSSSSLTGVSSSQPIQHPQQQGIQOTAPPOQTVQYLSQTSSEATTAQVPSPQ 777
 DB 719 NQGP-SSSLAGVLSQVQHP-QQGGIPVPPQAVQYSLPQAASSSEG-TVQPVSQ- 774
 QY 778 QAPVULPVQVSAQKO-STQGVSVQVAPAEVVAQPCATOPTLLASVDSASHDVASGMSDG 836
 DB 775 -----PVYSAQTQSTQGVSAQAPPEQPIQSQCTQVPLVSSVDSASHDVASGMSDG 828
 QY 837 NENVPSSSRHRCRTTKRHYKSVRSRSHKTSRKRLILVNSKGDVCEQLETHNR 896
 DB 829 NENAPSSSRHRCRTTKRHYKSVRSRSHKTSRKRLILVNSKGDVCEQLETHNR 888
 QY 897 KMTTFEFDLDGNDPEELATIMVNDLILIEREFVDYREIIEKADMLSECVSEYEPG 956
 DB 889 KMTTFEFDLDGNDPEELATIMVNDLILIEREFVQVREIIEKADMLSECVSEYEPG 948
 QY 957 DQGLJESLQKDYGFSGSQKLEGEFYQPIPASSMPQOIGITSSLTQVHSAQRPVSP 1016
 DB 949 DQGLJESLQKDYGFSGSQKLEGEFYQPIAVSSMPQOIGITSSLTQVHSAQRPVSP 1008
 QY 1017 VPESRLRESKVFPSHITDVTAASTQSPGMLSHSASLSLQAFSEILRAQMTREGPNTA 1076
 DB 1009 VPESRLRESKIFTESEIPDPVAASSTQSPGMLSHSASLSLQAFSEILRAQMTREGPNTA 1068
 QY 1077 PPMFHTGTFEPVPPPLSSIGAVPTTAANT--APVATSPNDISTSVIOSEVTPTE 1134
 DB 1069 PPMFHTGTFEPVPPPLSSIGAVPTTAANT--APVATSPNDISTSVIOSEVTPTE 1125
 QY 1135 EGIAGVATSTGVVTSGLPIPPVSESPVLSVSVSITIPAVSISTSPSLQVPTSTET 1194
 DB 1126 KGIQGVVTSQVAVASGGLTTLVSEETLSSAVSSSTAPAVVTVSTTSCQVQFTSGS-- 1183
 QY 1195 VVSTALYPSVTVSATSASAGSGTATGPVPPAVVSOQAAAGSTTVGATLTLSVSTTSPS 1254
 DB 1184 IASSTGSPGTFSTTGTTVVSVAVANAPKPTVTLQOVAVGNTAGVAVIVSVSTTPPFA 1243
 QY 1255 TASQLSLOLSSSTPLLAETVVVVAHSLDKTSHSSSTGGLAFSLAASSSSSPGAGVSY 1314
 DB 1244 MASQPSPLSSSTAPLAEVVAASLDAHSSSTAGLGLSCAPSSSSSTAVASS 1303
 QY 1315 ISQPGHPIVPIPSVIASTPLIPQAGPTSTPLIPVPSIPPLVQVAVNPVAVOQTLIHS 1374
 DB 1304 VSQGIHPIVPIPSVIASTPLIPQAGPTSTPLIPVPSIPPLVQVAVNPVAVOQTLIHS 1363
 QY 1375 QPQFALLPNOPTHCPPEVSDTOPKAPGIDDKTLEKLSLSEBSSSGAQAHSVLET 1434
 DB 1364 QPQFALLPNOPTHCPPEVSDTOPKAPGIDDKTLEKLSLSEBSSSGAQAHSVLET 1423
 QY 1435 SLVIESVTPGIPITTAVASKLLTSTSTCLPPTNLPLGTVALVPTPVVPGQVSTP-- 1491
 DB 1424 PLVVE-IVTGPITTAVASKLLTSTSTCLPPTNLPLGTVALVPTPVVPGQVSTP-- 1482
 QY 1492 --VSTTSGVKGPGTAPSKPPLTKAPVPLVGTTELPAITLPEQULPPEPSPSLTQSOELE 1548
 DB 1483 ASAPASTATGAKPGTTPPKPSLTKTVPVGTTELPAITLPEQULPPEPSPSLTQSOELE 1542
 QY 1549 DLDAQRLRTSLPEKITTSAVGPVMAAFTAITTEAGTQPGQGVQYKGPPLAASSGAGV 1608
 DB 1543 DLDAQRLRTSLPEKITTSAVGPVMAAFTAITTEAGTQPGQGVQYKGPPLAASSGAGV 1599
 QY 1609 FKMGFFQVSVADAQAGKXKXSEDAKSVHFESSSTSESVSSSSPESTLVKPEPNCITTI 1668
 DB 1600 VKMGFFQVSVADAQAGKXKXSEDAKSVHFESSSTSESVSSSSPESTLVKPEPNCITTI 1659
 QY 1669 FGISDVESAHKTTTASAKSDTQPTKVGRFOVTTTANKVGRFSVSKTDXITDTKXEG 1728
 DB 1660 SGISLDVDSHRTTTPKAKSETQPTKVGRFOVTTTANKVGRFSVSKTDXITDTKXEG 1719
 QY 1729 PVASPPMDLEQAVIPAVIPKKEKELSEPHLNKPPSSDPEAAFLSRVDQSSGPHPH 1788
 DB 1720 PVTS--PFPSDSEQTVIPAAIPKKEKELSEPHLNKPPSSDPEAAFLSRVDQSSGPHPH 1778

QY 1789 QLSKSLPSQNLQSLSSNSFNSSTYMSDNESEDIEDDLKLELRRLDKHLKEIQLDLSRQ 1848
 DB 1779 HLCSKSLPILQTLQSLSSNSFNSSTYMSDNESEDIEDDLKLELRRLDKHLKEIQLDLSRQ 1838
 QY 1849 KEIETSLTKLQKVPVAVIIPPAFLSGRRRPTSKSGSSRSSSLGKSPQLSGNTLSG 1908
 DB 1839 KEIETSLTKLQKVPVAVIIPPAFLSGRRRPTSKSGSSRSSSLGKSPQLSGNTLSG 1898
 QY 1909 QGAASVLPQQLHPHPGIPESQNLQPLXPSPSSDNLVSAFTSDGAISSPSLSAPQ 1968
 DB 1899 QGSTVLNQLQTLHPHPGIPESQNLQPLXPSPSSDNLVSAFTSDGAISSPSLSAPQ 1958
 QY 1969 GTSSTNTGATVNSQAQAQAPPAATSSKRGTTTDLHLKLVMDWARDAMTLSSRGSKGM 2028
 DB 1959 GTSSTNTGATVNSQAQAQAPPAATSSKRGTTTDLHLKLVMDWARDAMTLSSRGSKGM 2018
 QY 2029 NTEGPMARKFAPQGLCTSMNSNGSPISAASATSLGHFTKSCPPQOYGFPAATPG 2088
 DB 2019 NTEGPMARKFAPQGLCTSMNSNGSPISAASATSLGHFTKSCPPQOYGFPAATPG 2078
 QY 2089 AQMSGTGAPQPLQGFQVGTASLQNFNISNLQKSIINPSPSNLRTT 2136
 DB 2079 TQMSGTGAPQPLQGFQVGTASLQNFNISNLQKSIINPSPSNLRTT 2126
 RESULT 4
 Q9P1S9
 ID Q9P1S9 PRELIMINARY; PRT; 670 AA.
 AC Q9P1S9;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Kinase deficient protein KDP (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC
 RX MEDLINE=20127920; PubMed=10660600;
 RA Moore T.M., Gary R., Johnson C., Coptrout M.J., Ridley A.J.,
 RA Morris J.D.H.;
 RT "PSK, a novel STE20-like kinase derived from prostatic carcinoma that
 RT activates the JNK MAPK pathway and regulates actin cytoskeletal
 RT organization.";
 RL J. Biol. Chem. 275:4311-4322(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DB EMBL; AF061944; AAF31483.1; -;
 DB GO; GO:0005524; F:ATP binding; NAS.
 DB GO; GO:0004674; F:protein serine/threonine kinase activity; NAS.
 DB GO; GO:0006468; F:protein amino acid phosphorylation; NAS.
 DB InterPro; IPR006719; Prot kinase.
 DB InterPro; IPR008271; Ser_Thr_kinase.
 DB Pfam; PF00069; Pkinase; 1.
 DB ProDom; PD000001; Prot_kinase; 1.
 DB PROSITE; PS00101; PROTEIN KINASE DOM; 1.
 DB PROSITE; PS0108; PROTEIN KINASE ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 FT NON TER 670
 SQ SEQUENCE 670 AA; 73695 MW; C24A9A8D7FC32AA CRC64;
 Query Match 31.5%; Score 3408; DB 4; Length 670;
 Best Local Similarity 100.0%; Pred. No. 6.1e-159;
 Matches 668; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSGAAEKQSTPGSLPSPPAPKXGSSDSSVGEGLGAADAAYVGRTEEXRRRHT 60
 DB 1 MSGAAEKQSTPGSLPSPPAPKXGSSDSSVGEGLGAADAAYVGRTEEXRRRHT 60
 QY 61 MDKDSRGAATTTTTHRRFFRRSVICDSNATALELPGLELPPPSIPAAVQSAAPPEPH 120
 DB 61 MDKDSRGAATTTTTHRRFFRRSVICDSNATALELPGLELPPPSIPAAVQSAAPPEPH 120

DR GO: 00004713; F:protein-tyrosine kinase activity; IEA.
DR GO: 00006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR000719; Prot. kinase.
DR InterPro: IPR002290; Ser thr pkinase.
DR InterPro: IPR008271; Ser thr pkin AS.
DR InterPro: IPR001245; Tyr_pkinase.
DR ProDom: PD000001; Prot. kinase; 1.
DR SMART: SMO0220; S_TKc; 1.
DR SMART: SMO0219; TyrKc; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 2219 AA; 234470 MW; B7AD70196C2A92C0 CRC64;

Query Match	24.2%;	Score 2615;	DB 4;	Length 2219;
Best Local Similarity	31.9%;	Pred. No. 2.4e-119;		
Matches 811;	Conservative 234;	Mismatches 645;	Indels 856;	Gaps 86;

QY	60	TMKDSRGA-----	AA	TTTTTTRFRFRSVICDSNATALELPGIPLSLPQSP	IPAA	110	
Db	8	TLMEPGGAPGAPMAEBRPAARAPGPRFLRRSV--	ESDQ	EEPPGLR-----	AAE	56	
QY	111	VPOSAPPEPPEEREET-----	VTAA	TQSVVAQGPAA-----	AAGQAVAGPASTV	156	
Db	57	APBQGPQPLQGRVLLLCTRRLLAERARGRPAAPDAALVAOPGAGADAGPEP---			113		
QY	157	PSSTKDRPVQPSLVGSKKEPP-----	ARSGSGG	GSAAK-EP-----	QOE	RSQ 139	
Db	114	---VGTQEPDPDIAAVETAPAPADGGPRREAAATYKDEGGAEAKEPFGRRRRDPEE			170		
QY	200	QODDIEELEFKAVGMSNDGFLKFDLIEIGSFRTYVKGSLDTETTYEVAWCELODRLT			259		
Db	171	EEDEDEDL--KAVATSLDGRFLFDLIEIGSFRTYVKGSLDTETTYEVAWCELODRLT			228		
QY	260	SEORFKEAEAMKGGQHPPIVAFYOSWSTVYGKXCIVLYEELMNSGTLKTLKPKVM			319		
Db	229	LERQKKEAEAMKGGQHPPIVAFYDPMWSSAAGKXCIVLYEELMNSGTLKTLKPKVM			288		
QY	320	KIKVLRSWCROILLKGLQPLHTRTPPIHRDLKCDNFIPTGTSVKGIDGLATLRASF			379		
Db	289	KEKVLRSWCROILLKGLFLHTRRPIIHNDLKCDNFIPTGTSVKGIDGLATLRASF			348		
QY	380	AASVIGTPEFMAPEMYEKEYDSVDVYAFQCMLEMATSEYPYSECQNAQIYRRTYSGV			439		
Db	349	AASVIGTPEFMAPEMYEKEYDSVDVYAFQCMLEMATSEYPYSECQNAQIYRRTYSGI			408		
QY	440	KPASFDKVAIPYKKEIIEGTRONKDERSYDOLLNHAFFQETGRVLAEBDDEK--			497		
Db	409	KPASFEKHDPKELIIGECICQNKEREIKDLSHAFPAETGRVLAEBDHGRKST			468		
QY	498	IAIKMLRIEDIKUKGKYKDNALIEFSDLERDVEDYAOEKNVESGYVCEGHHKTMAA			557		
Db	469	IALRIW--VEDPKKUKGKPKONGALIEFTDLKETPRDVAQEMIBSGFFHESDVKVAAS			526		
QY	558	IKDRSLIKRAKEQOROLVBEQEKKQOEBSSIKQOEVGSSASQTIKOLPSASTGIPTAS			617		
Db	527	IKDRVALLQWRE--RIWPALQPKQEDVGS-----	PDKAR	GPVPL	566		
QY	618	TTSASVSTQV-----	EPEEP	ADHQOLQYQO	PSI-----	SVLSNGTYDSQGS	SVPE 666
Db	567	QVOYVYHQAQAGPGPEPEPEPADQH---	LIP	PLTISATSLASDSTFDSQGS	QSVTS	622	
QY	667	SRVSSQQTIVSYGSGHQAHS TG-----	TVPGH	IPE	TVAQSGPHG-----	706	
Db	623	SG--SSQSGVMGLADAPASPAGQVCSPVSGPYLPGSLPS--	LGA	YQGP	TAAPGLPVGS	680	
QY	707	-----	VP	PSVAQSGSGQSPS-----	723		
Db	681	VPAPACPSLQOHPEDDPAMSFAPVLPPESTWPTBPQGPAPGQGPPLAQCTPLPQVLA			740		
QY	724	-----	SSSL	UGVSSQPIQH	QOO--QSIQO-----	747	

Dh	741	POVVVVLQGVNPHRLPRVYLTAASQVGA	RAQDLQ	QMLQ	QALQ	DLQ	ALQ	VLQ	VP	MP	PL	VV	PP	IT	800
Qy	748	-----	TAP	QOQ	TVQ	Y	-----	-----	-----	-----	-----	-----	-----	757	
Dh	801	PLAGIDGLP	PL	DL	ET	AT	VP	VP	PP	QV	SP	PA	VI	PL	860
Qy	758	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	757	
Dh	861	GAPL	MP	CK	TI	VN	NA	PA	TI	PL	LA	VA	PC	VA	920
Qy	758	-SLSQ	T	S	T	S	S	E	A	T	T	A	Q	P	920
Dh	921	PSPH	TV	Q	M	K	A	T	PP	Q	AL	PP	Q	PT	980
Qy	779	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	778	
Dh	981	LPQ	V	L	PP	Q	A	L	P	V	R	P	R	E	1040
Qy	779	-----	A	P	V	L	P	Q	-----	V	S	A	K	Q	802
Dh	1041	PA	V	L	P	P	L	E	V	L	P	A	B	E	11000
Qy	803	-----	E	P	V	A	Q	P	A	Q	P	T	L	A	849
Dh	1101	SP	C	P	T	V	Q	L	T	V	P	V	E	Q	1159
Qy	850	RT	T	K	H	R	K	S	R	H	E	K	T	R	909
Dh	1160	RA	A	R	H	H	R	S	T	R	A	S	R	O	1219
Qy	910	P	E	B	A	T	I	V	N	N	D	I	L	A	965
Dh	1220	P	E	B	A	T	I	V	N	N	D	I	L	A	1279
Qy	966	K	D	O	G	F	S	G	S	O	K	L	E	G	1025
Dh	1280	L	G	-----	T	C	E	B	R	Q	Q	A	N	A	1321
Qy	1026	K	V	P	S	E	I	T	D	V	A	S	T	A	1085
Dh	1322	-----	P	E	S	S	P	L	P	L	S	L	P	E	1368
Qy	1086	T	F	P	V	P	P	L	S	I	A	G	V	P	1144
Dh	1369	S	N	P	-----	P	G	A	P	-----	P	A	P	A	1400
Qy	1145	G	V	T	S	G	L	P	I	P	V	S	E	P	1204
Dh	1401	-----	M	P	B	-----	-----	-----	-----	-----	-----	-----	-----	1404	
Qy	1205	V	T	V	A	T	S	A	S	A	G	S	T	A	1264
Dh	1405	-----	A	S	G	T	S	O	A	G	-----	P	G	-----	1449
Qy	1265	S	S	T	P	L	A	E	V	V	A	S	A	L	1323
Dh	1450	P	C	I	P	A	P	A	A	S	T	R	O	A	1496
Qy	1324	L	V	I	P	S	V	A	S	T	P	L	P	Q	1383
Dh	1497	A	A	V	A	N	S	L	A	T	S	O	L	P	1526
Qy	1384	Q	P	H	N	C	E	V	N	D	O	P	A	R	1440
Dh	1527	-----	S	D	G	B	O	P	P	P	V	G	V	D	1573
Qy	1441	T	V	P	G	I	P	T	V	A	S	K	L	I	1499
Dh	1574	-----	G	R	D	F	L	I	B	R	G	Q	R	S	1614
Qy	1500	K	B	T	A	S	K	P	R	L	T	A	P	V	1556
Dh	1615	-----	O	P	I	V	E	S	K	E	L	A	P	T	1657

QY 1557 TLSPKIVTSAVGPVMAPTAITEAGTOP-----QKGVQVKEGPV-LATSSGAGVF 1609
 Db 1658 VASDSHV-----VPSVPQVPAVPAVPAVPELPDRDROGEAGSAPPEPDMGTVGQASH 1712
 QY 1610 KMGKPCVSAVDAQAQKQKNGKSDAKSVHESSTSSSVLSSSSPESTLVKPEPNGITTP 1669
 Db 1713 PQTIGRALGSPRKREQQDVSSPATVGRFVSIVTQDEWTLASPHSLKSAAPPVYLD 1772
 QY 1670 GISS-DVPESAHKT-TASEAKSDTGOPTVYGRFOVTTANKVGRSVSTEDKIDTKE 1727
 Db 1773 AESSPQVLAVRRAQTAASSIEVGVGEVPS-----SDSGDE 1807
 QY 1728 GVAAPPEPMDLQAVLPVATPKKKEPSEPHLNGSSDPPEAPLSPDVDDSGSPHP 1787
 Db 1808 GPAPRPVVO--KQASLPV-----SGSVAGDPYKATATLQPSRAGSLGEPF 1853
 QY 1788 HOLSSKSLPSQNLSSQSLNSFNSSYMSDNESEDIEDEDLKELRLRDLRHLKEIDLOQR 1847
 Db 1854 SRVGMK-VPTISTVFSH--QSSYISDNDSELEBADIKKELQSLREKHLKESLQSO 1909
 QY 1848 QKHEISLYTKLGK-VPAVITLPPAALSGRRRPPTKSGKSSRSSSLGNSSPOL---S 1903
 Db 1910 QKQELBALVRLGKPLPNNVGFHTAPPTGRRRKTKSKLAKGKLLNPLVRLQKVAASST 1969
 QY 1904 GNLSSGSAASVLRPQOTLHPGNTPE-----SGQN-OLLQD--LKPSPSSDN 1947
 Db 1970 GHILADSSRG-----PPAKDPQOASVGLTADSTGLSGKAVOTQOPCSVNASLSD 2018
 QY 1948 LVSAPFSDGALVSPSLAPGQ-----TSSINTV----- 1976
 Db 2019 ICSGLASDGS-----GARGQWTVVHPHSERTVYSSSKPRARPLSGPVSIVSIALKR 2072
 QY 1977 -----GATNSQAQAQP---PAM-----TSSKGTFTDHLQVNNMADAM 2016
 Db 2073 ICLGKEHSSRSSTSLAPGPEPPOPALHYQAQVNNNNKKGFTDHLVEMWTSKTV 2132
 QY 2017 NLSGRRS-----KGMHNVGSGMARKSAPQOLCISMTSNLGSASPISA-- 2061
 Db 2133 GAALQKFTLNQLKOTOKLQDMEAOAGAPABGAPAMTAP--RAGVGMFRLPPAPGFLSTTV 2191
 QY 2062 --ASATSLGHFTKSMCPQOYGPAT 2085
 Db 2192 IPGAAPTLVPTPGSGCPRAVSTPTS 2217
 RESULT 7
 Q8TCX6 PRELIMINARY; PRT; 1800 AA.
 AC Q8TCX6;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN PRKMNK3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Holden S.T., Raymond F.L.,
 RT "Full length coding sequence for human MNK3 determined from Clontech
 Maratlon cDNA (fetal brain)."
 CC Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AY082340; AAL99253.1;
 DR Genew; HGNK:14543; PRKMNK3;
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR000719; Prot. Kinase.
 DR InterPro; IPR008271; Ser. Thr. Pkin. AS.
 DR Pfam; PF00069; pkinase; 1.
 DR Prodom; PD000001; Prot. kinase; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE-ST; 1.
 KW Hypothetical protein; ATP-binding; Kinase;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 1800 AA; 198442 MW; 9C093A9F0BAC2117 CRC64;
 Query Match 22.8%; Score 2463.5; DB 4; Length 1800;
 Best Local Similarity 33.2%; Pred. No. 47e-112; Indels 621; Gaps 79;
 Matches 750; Conservative 253; Mismatches 635;
 QY 25 PKNGSSSDSSVGEKLGAAADAVTGRTEYRRRRHTMDKDSRGAATTTTTRRRFRSGV 84
 Db 8 PANTEDSEKPDISTENRNPQVAAITLYEARLK-----ENKSTFMSAGTEVERKRRFRKSV 63
 QY 85 ICDSNATALELPLSLPQPSIPAAVPGSAPPEPHREFTVATATSGVAAQPPAAAAG 144
 Db 64 -----EMTEDDKVAESSP-----KDERIKAAANNIPRVDLPSNVLRG 100
 QY 145 EQAVAGPASTVPSSTSKDRPVSQPSLVSKKEPPRARGSGGSAKEPQERSSQGDID 204
 Db 101 GQEVKYEQCSKSTSEISKD-----CFKXNKEKME----- 130
 QY 205 EELETAYAGMSNDGRFLKFDIEIGRSFKTVYKGLDTEETVEVAMCELODRKLTKEER 264
 Db 131 EEAEKMAKAVATSPSGRLKNDIELGRGAFKTVYKGLDTEVWEVAMCELODRKLTKEER 190
 QY 265 FKEEAEMLKGLQHPNIVRFYDSWESVKGKCLIVTEIMTSGLTKYKRFVMMKIKYL 324
 Db 191 FKEEAEMLKGLQHPNIVRFYDSWESILKGGKCLIVTEIMTSGLTKYKRFVMMKIKYL 250
 QY 325 RSWCRQILKGLQHTRTPEIIRHDLKCNIFITPTGSKIGDLGLATLKRAFPKSVI 384
 Db 251 RSWCRQILKGLQHTRTPEIIRHDLKCNIFITPTGSKIGDLGLATLKRAFPKSVI 310
 QY 385 GTFEPFAPPEVKEEYDESVDVYAFGCMLEMATSEYPSYSCQAAQIYRRVTSGVKPRSF 444
 Db 311 GTFEPFAPPEVKEEYDESVDVYAFGCMLEMATSEYPSYSCQAAQIYRRVTSGVKPRSF 370
 QY 445 DKVAIPVEVKEITGICIRONKDERYSTKDLINAFPOEFTGVVLEAEED--GEKTAIKL 502
 Db 371 NKVTDEPVKEIIEGCIROKSESLIRDLINAFPAEDTGLRVLEAEEDCSNSLRL 430
 QY 503 WLRIEDIKLKKGXKNNEALEFSPLEEVPEVPAQEMVSGVCCGDHKTMAKAIKRV 562
 Db 431 W--VEDPKLKKGXKNNEALEFSPLEETPPEVPAEMVSGVGFHESDSKAVAKSIRDRV 488
 QY 563 SLIKRREQRQLVREDEKKEKKESSILKQVEOSSAQTGIKQLPSASTGIPASTTSAS 622
 Db 489 TPIKKTRE-----KPRAGCLIEHRDSQCKSMGVPPQPQ--NTLLPLA---PA 531
 QY 623 VSTQVEEPEEADQH--QQLQYQPS--ISVLSGTVDSSGSSVFTERVSSQYVSY 677
 Db 532 QQTGABCEETFEVQDHYVQQLLRKPOCHSCSVTGDNLSFGAASVHISDT--SSQPSVAY 589
 QY 678 GSGHEQHSNGTVPGHIPSTV--QAQ--SQPHGVPPSSVAAQSGQSSSSSLTVSSSP 735
 Db 590 SSNQ-----TWGQVNSINIPQAEVNVPPQIT-----SSQQL 620
 QY 736 IQHPQOQGGIQOTAPPOQTVOYSLSTSSSEATTAQPVSPQAPQVLPVYSAGKSGSTG 795
 Db 621 VGHQGVVSGIQKSKLQPOLPLVYQO-----STVLPV--HVLGPTVVSQPVSPPLTVQK 674
 QY 796 VSGVAP--AEVPAVAQPOATOPTTASSVDASHVYASGMDGNNTV-----PSSSGHGG 849
 Db 675 VPQIKPVSPVGAQQAALLKPDV--VRSINQDVAT--TKENAVSSPDNPSGNGKQDR 727
 QY 850 RTTGRHRTKSVRSRREKTSRPLRTILNNSKDRVVECOLETHNKKMTTFKFDLDGDN 909
 Db 728 IKQR-----ASCPRBKGTQTLVQVSTSGDNVVECOLETHNKKMTTFKFDVGDGA 781


```

QY 910 PEEATINWNNDFILAIERESFVDQVREIIEKADMLSEDVSEVEDEGQJESLQKXDY 969
DB 782 PEDIALDVWENDNFVLESEKEFEVELRAIVGQAQIL--HHAFERATQVDISTIVDSNS 839
QY 970 GFSGSQKLEGEKQCIIPASSMPOQIGIPTSSLTQVHAGRRFIVSPVESRLRESKVPF 1029
DB 840 SQTSS-----EVOVINSTSTQTSNESAPQS--SPVGRWRF-- 874
QY 1030 SEITDTVAASTAQSPGMNLSHSSASLSLQQAFFELRRQOMTEGPRVATAPNSHGTPTFPV 1089
DB 875 --INQTIENRETQS-----PPSLQH----- 892
QY 1090 VPPPLSLIAGVPTTAATAAPVATSPPN--DISTVISOEVTPTEEGIAGVATSTGVYT 1148
DB 893 -----SMAVAP-----GRHPLPSPKNTSNKEISNDTLTINNCHRALFTFSKSHKDV 942
QY 1149 SGGPIRPVSESPVLSVSVSSITIPAVVISITTSPSLQVPTSTSEIV--VSSTALYPSV 1205
DB 943 DG-----KISECAVETKOPALIVQVEDNRQJMAVPTNSSYSTT 982
QY 1206 TVSATASAGGSTATPGPKPRAVVSQQAAGSTTGALTSTSTSTSTSTSTSTSTSTSTST 1265
DB 983 SVRAVPAECBELTQASIFITVPCHQ-----TASQADALMSHPESTQ----- 1026
QY 1266 STSTPTLAETVAVSAHSLDKTSHSSTTGAFSLAPSSSSPGAGVSSYSQPGALHPLV 1325
DB 1027 -----TSGNSLTTLIADQKQPTLSVQQPMADAFISIQE----- 1060
QY 1326 IPSVIATSTPLPQAAGSTPTPLPQVPSIPPLVVPVAVVAVQOTLIHQOPALPLNQP 1385
DB 1061 -ETVNTBASSPKVITPQTGGE-----PTTLP----- 1089
QY 1386 HTNCPVEDSDTOPKAPGID--IKTLEKLS--LFEHSSSG--AQHASVLETSVIE 1439
DB 1090 -TTLVLEDGRPRPLEFADNRKIKLDEKLNLYQHSISISIPESKOTQSDSPSS 1148
QY 1440 STYTPGIPPTAV-----ABSKLLTSTTCLPTNPL----- 1472
DB 1149 AEDTLSCPTVEVIAISHGKIKDSFVQSPNFQJGSKLNSVAA--QPANISVFKRLNV 1206
QY 1473 -----GYALPVTVPVTPGQVST--PVSTTSGVYKPG--TAPG--KEPLTK 1512
DB 1207 ITVSPELCHHEMSDASLPGDEAVPAVSSGAIHLQGGGFGISFTCPSLKMPISK 1266
QY 1513 APVLPVGTBLPAGTLSEQLPPPPSPSLTOSQOPLDLQALRRTS-----PEX 1563
DB 1267 -----KSWRK--LKSMAVRLQOSTSFFKRSVROVETE 1298
QY 1564 TVTSANGVPVSMAPTAITEAGTOPKGVQKGVKGPVLTATSSGAGVFMGRFOVVAADGA 1623
DB 1299 EMMSALAP--DPIPLTRSTADTRALNRCK-----AMGSPQORGFQYITTPQ-- 1344
QY 1624 OKRGKKSEDAKVHFESSSTSESYSVS--SSPSTLYKPEPNITTPGLSSDVPEBAHKT 1682
DB 1345 QQSARKTSGEIEHTSVFSETNHSSEAFIKTAASQVLETP----- 1385
QY 1683 TASEAKSDTGQPTKVGAFVTTANKVGFVSXKTEDKITDTKEG--PVASPPMDLE 1739
DB 1386 -----ATQNKTSFSYKLO--AQETKEKGVYKQGDNPLSFS 1423
QY 1740 QAV--LPVITPKKEPE-----LSEPSHLNG--PSSDPE--AAFLSRDV 1777
DB 1424 AACETDVSVTPEKEFEFTSATSSMOSGSEILLKREILITAGKQSSDSFBSASIA-- 1480
QY 1778 DDGSG-----SPHS--PHQSSKSLJSONSGLSNFS--NSVYMSDNESDIEDD 1825
DB 1481 --GSGKSVAKTGPESNOCLPHH-----EQAVYAOQSLLFYSPSPMSDDESEIIEDD 1532
QY 1826 LKLELRRLDKHLKEIODLOSROKHIEBSLYTKLGVPVPA-----VILPPAAPLSGRRR 1879
DB 1533 LKVELQRLREKRIQVAVNLTQONKELQELYERLKSIRKQSTQSTIEIPLPASP-----R 1587

```

```

QY 1880 RPTSKSKSSRSSSLGNKSPOLSGNLSGSAASVILHQOTILHPGNIPEGQNLQPL 1939
DB 1588 RPSFKSKLRSR-----FQSLTHVDNGIIVAGKSLINEL 1622
QY 1940 KPSPSDNLYAFTSDGAI SVPSLSAPGQSTSTNTVATVANSQAQAQAPAMISSRGT 1993
DB 1623 E-----NPLCVESNMAASCCQSP--ASKRM 1645
QY 2000 FTDDLHLKLVNDWADAM-----NLGRRGSKGHMNYEGGMARKEFAPQL-----CI 2047
DB 1646 FTDDLHLKLVNDWMTKEAVGNLSIKPSLNLQKQSKHLENNKXSE--NTPSTMGTSWTI 1704
QY 2048 SMTNLSGSAPIASAATSLGHFTKSMCPPOQYGP-----APFGAQMSGTGPA 2098
DB 1705 SLSLQIRGAFTSLPQGLSLPSFPQ-----PLSSYGMFHVQYNAVAGAGYPQWVGISGT 1761
QY 2099 POPU-----GQFQPVGTASLQNFNTSNLQKSIJNPG 2130
DB 1762 QQSVVIPAQSGGFPQ--GMNMQAFPTTSVQNPATITPG 1798

```

RESULT 8

```

ID Q9BYP7 PRELIMINARY; PRT; 1743 AA.
AC Q9BYP7;
AD 01-JUN-2001 (Tremblrel, 17, Created)
DT 01-JUN-2002 (Tremblrel, 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel, 25, Last annotation update)
DE Protein kinase MNK3.
GN PRKNK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21455683; PubMed=11571656;
RA Verissimo F., Jordan P.;
RT "MNK kinases, a novel protein kinase subfamily in multi-cellular
RT organisms."
RL Oncogene 20:5562-5567 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Jordan P.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ409088; CAC32455.2; -
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferrase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr_pkin_AS.
DR Pfam; PF00069; pkinase; I.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_SF; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1743 AA; 191788 MW; 7B06C8C7B8E4D57 CRC64;

```

Query Match 22.6%; Score 2439; DB 4; Length 1743; Best Local Similarity 33.6%; Pred. No. 7.1e-111; Matches 746; Conservative 246; Mismatches 617; Indels 610; Gaps 80;

```

QY 27 NGSDDSVGEKLGAAADAVGTGTEERYRRRHNTMDKSRGAATTTTTEHFFRRSYIC 86
DB 18 DGISEFNRVPO--VAATITVARKLE-----KSTFSASGETVRRKFFKSV-- 63
QY 87 DSMATALRLPGLPLSLPQPSIPAAVPOSAPPEPHREFTVATATQVAAQPPAAAPGQ 146
DB 64 -----EMTEDDKVAESSP-----KDERIKKAMNIPRDYKLPNSVLRGQ 102

```

QY 147 AVAGPAPSTVPSSTKDRPVSOPLVSGXKEEPPASGSGGSAKEPOEERSQOQDDIEE 206
 DB 103 EKVYECOSSTISISD-----CFKKNKEME-----BE 132
 QY 207 LETKAVGMSNDGRFLKEDIEIGRGSFKYKGLDTEITVEVAMCLODRKLTKSERORFK 266
 DB 133 AEMKAVATSPSGRFLKEDIEIGRGAFTYKGLDTEITVEVAMCLODRKLTKSERORFK 192
 QY 267 EEMKMLKLOHNPVYFYSWESTVWGKCTIVYTELMTSGTLTKLXPKMKIVLVS 326
 DB 193 EEMKMLKLOHNPVYFYSWESTVWGKCTIVYTELMTSGTLTKLXPKMKIVLVS 252
 QY 327 WCAQILKLOHNPVYFYSWESTVWGKCTIVYTELMTSGTLTKLXPKMKIVLVS 386
 DB 253 WCAQILKLOHNPVYFYSWESTVWGKCTIVYTELMTSGTLTKLXPKMKIVLVS 312
 QY 387 PERMAEMVEEKYDESVDVYAFGMCMLEMTSEYVSECQNAQIYRVTSGVKPASFDK 446
 DB 313 PERMAEMVEEKYDESVDVYAFGMCMLEMTSEYVSECQNAQIYRVTSGVKPASFDK 372
 QY 447 VAIPVEKTEIEGCRONKDRYSIKOLINAFEOELGAVVELAEEDD--GEKIAIKML 504
 DB 373 VTDPEVEKEIEGCRONKDRYSIKOLINAFEOELGAVVELAEEDD--GEKIAIKML 431
 QY 505 RIEDIKLKXKYNKDNEMIESFPLERDPEDVACQEMVSGVCEGDHKTAKAKIKRVS 564
 DB 432 -VEDPKLKXKYNKDNEMIESFPLERDPEDVACQEMVSGVCEGDHKTAKAKIKRVS 490
 QY 565 IKRREOROLVREOEKKEESLKOQVOSASQGTIKOLPSASTGIPASTTSASVS 624
 DB 491 IKRREOROLVREOEKKEESLKOQVOSASQGTIKOLPSASTGIPASTTSASVS 533
 QY 625 TOVEPEPEADQ--OOLYOQOPS--ISVLSGTVDVSQOQSSVFTSRVSSQOQTVS 679
 DB 534 TGACEETEVDOHVRQOOLQORRPOQHCSVYGDMLSEAGASVYHSDT--SSQSVASVS 591
 QY 680 QHEOAHSTGVYHPISTV--OAO--SOPHGVYPPSSVAVQOQOQSSSILTVSSQOPIQ 737
 DB 592 NO-----TWGSOVSNIPQAEVNVPOQIT-----SSQOQV 622
 QY 738 HPQOQOQIOQTAPRQOQTVOSLSQTSSTSEATTAQVSPQAPQVLPQVSAKQSTQVS 797
 DB 623 HYQVSGIKKXSKLTQOQILPLVQOQ--STVLPLV--HVLGPVVSQPVSPILTVQVP 676
 QY 798 QVAP--ABPVAVAQOQATOPTLLASSVDSASDVSAGSDGEMV-----PSSGRHGR 851
 DB 677 QIKVVSQPVGAEOQOALIKPLD--VRSINQDVAT--TKENVSPPDNPSGNGKQORIK 729
 QY 852 TKRHYKRSVRSRHEKTSRPKLRIINYSNKGDRVVEQOLETHNRKMYTFFEDLDGNPE 911
 DB 730 QRR-----ASCPRPEKGTFOQLTVLOVSTSGDNVVEQOLETHNRKMYTFFEDLDGNPE 783
 QY 912 EIAITMVNNDPILAIERESFVDOVREIEKADMLSEDSVVEPBGDQLESLOKQDYGF 971
 DB 784 DIADVMEDNVLSEKKEKFEELRAIVGQOQEL--HVHATERATGVDSITVDSNSQ 841
 QY 972 SGOQKLEGEFKOPIPASSMPOQIGIPSSLTQVHASAKRRTVSPVPSKRLREKVPSE 1031
 DB 842 TGSS-----EOVQINSTSTQTSNESAPQ--SPVGRMREC-----874
 QY 1032 ITDITVASTQSPGNLHSHASLSLQAFSELRAQCTEGPNTAPPEFSTGTGFPVVP 1091
 DB 875 INQITRNKETS-----PESLOH-----892
 QY 1092 PFLSIAQVPTTAATAVAPATSSPPN--DISTVLOSEVTVPTBEGIAVASTGVTS 1150
 DB 893 ---SMSAVP---GRHPLSPKNTSNKEISMDTLTIENNPCHRALFTSSSEHDDVVDG 944
 QY 1151 GUPIPVSESPLVLSVSSITIPAVVSIISTSPSIQVPTSISEIY--VSSITALYPSVT 1207
 DB 945 -----KISECASVETKOPAILYOVEDNRQIMAPVNTSSSYSTTSV 984

QY 1208 SATSASAGSTXTPGPKPPAVVSOQAGSTVATLTSVTTTSPSTASQLSQLSST 1267
 DB 985 RAVPAECBGLTKQASIFIPVPCQO-----TASQADALMSHGEESTQ-----1026
 QY 1268 STPTLAETVVAHSLDKTSHSTTGIAFSLAPSSSSPGAGVSSYISQPGIADPLV 1327
 DB 1027 -----TSGNSLTLTADQKRPQITSVQCPAMDFAFIQBG-----E 1061
 QY 1328 SVIASTPILPOAAGPTSTPLPOVPSIPPLVQPVANVAVQOTLHISQOPALLPNOPT 1387
 DB 1062 TVVNTIASSPKVIVIPQTPGLE-----PTTLQP-----T 1090
 QY 1388 HCEPVDSDTOPRACPID--IKTLEKLR--LSEHSSSGAQHASVSLSTVIESTV 1445
 DB 1091 TVLESDEGRPPKLEFADNRKTLDEKRLMLVQEHIS--STYPSQ-----1135
 QY 1446 IPTTAVAPKLTLSSTSTPLPNTNPLGVALPVTPVTV--PGVSTPST-----TTS 1497
 DB 1136 -----KOTSIDSPSSSADTLSCVTEVIAISHCGIKDSFYQSPNPOQTS 1183
 QY 1498 GVKPGTAPSKP-----LTKAPVLPVGTLEPAGTLPS--QLP--PEPGLTOS 1543
 DB 1184 KLSNVAASQPANISVFKRDLNVTSV-----SELCHMSDASLPGDPAYPAVVS 1238
 QY 1544 QOPELDLDAQRLRTLSPEKXITVTSANGVYSMAAPTAITEAGTOPQKQVSGVKEGPVLATS 1603
 DB 1239 -----GGNHLQTVETEMKSAIAP--DPIELRESTDOTALNCK-----1279
 QY 1604 SGAGVFKMGFPQVSVAAADQKQKSEDAKVHFESSSTSESSVLS--SSPSTLVKPE 1662
 DB 1280 AMSGSFORGRFOYITTPQ--QCSAKTSPGIEHITSVSEINHSSEAFITAKSOLVIE 1337
 QY 1663 PNGITIPGISSDVPESAHKTJASBAKSDTQPTKVGAFVYTTANKVGFVSKEDEKIT 1722
 DB 1338 P-----ATQNPKTSFSEKLO--ALQ 1356
 QY 1723 DTKKGG--DVASPPMDLEQAV--LDPAVTPKKEPE-----LSESHL 1761
 DB 1357 ETCKENKGVKQGDNPLFSNACETDVSVTPREKEFEITSATSSMOSGELLKKEEIL 1416
 QY 1762 NG--PSSDE--AAFLSRVDVDSG-----SPHS-----PHOJSKSLPSONLSQSLNS 1807
 DB 1417 TAGKOPSSDEFSASLA--GSGKSVAKTGPESNOCLBHH-----EODAYAQTOSSL 1465
 QY 1808 F--NSSYMSDNEDIEDLKLRLRDKHLKEIDLOSROKHEIESLYTKLGVP 1865
 DB 1466 FYSFSSPMSSDDSEIEDLKLRLRDKHLKEIDLOSROKHEIESLYTKLGVP 1525
 QY 1866 ----VITPPADLSGRRRPTKSKGSKSSRSSLSGKSPOLSGNLSGQSAASVLAHQ 1919
 DB 1526 KTQSTEIPLPPASP--RRPSFKSKLRSR-----PQS 1555
 QY 1920 TILHPGNIPESGQOQLQPLKPSPPSNDLVSAFTSDAISVPSLASQGTSTNTVGAT 1979
 DB 1556 LTH-----VDN-----GIVALDPLQVE 1572
 QY 1980 VNSQAQAOPPAMTSRKGFTDDIKLVNWARDAM-----NLSGRGSKGHMVEG 2032
 DB 1573 SNAASCOOSP--ASKKGMFTDLHLKLVMDWTEAVGNSLIXPSLQOLKOSQKLETEN 1628
 QY 2033 PGMAKTSAPGOL-----CISMTSNLGSAPISAAASATSLGHFTKSKCPPOQYGF 2083
 DB 1629 WNKVSE--NTPTSMVYTSWISLSLQIRGAVPTSPGUISLPSFG--PLSSYMPHYVCQ 1684
 QY 2084 ----ATPFGQWNGSTGAPAPOL-----GOFQPVGTASLQWENISNLQKSISSNPG 2130
 DB 1685 YNAVAGAGYVQWNGISCTTOQSVITPAQSGGPPQF--GMMQAPPTSSVGNPANTIPFG 1741

RESULT 9
 Q96J92 PRELIMINARY: PRT: 1243 AA.
 ID Q96J92
 AC Q96J92; Q96DT8;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative protein kinase MNK4 (Serine/threonine protein kinase)
 DE (EC 2.7.1.37).
 GN PRKMNK4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Christakov D.A.;
 RL Thesis (2001), Department of INSERM U36, College de France, Paris, France.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF390018; AAK91995.1; -
 DR EMBL; AJ316534; CAC48387.1; -
 DR Genew; HGNC:14544; PRKMNK4.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004651; F:GMP-dependent protein kinase activity; IEA.
 DR GO; GO:0004662; F:Protein kinase CK2 activity; IEA.
 DR GO; GO:0004674; F:Protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PSS0108; PROTEIN KINASE_ST; 1.
 KM Hypothetical protein; ATP-binding; Kinase;
 KW Serine/threonine-protein kinase; transferase.
 SQ SSQUENCE 1243 AA; 134739 MW; BAC35F0938BA391 CRC64;
 Query Match 19.7%; Score 2126; DB 4; Length 1243;
 Best Local Similarity 31.9%; Pred. No. 9.8e-96;
 Matches 612; Conservative 191; Mismatches 393; Indels 724; Gaps 58;
 QY 10 SSTFGSLFSPAPAPKNGSSDSVGEKLGAAADAVTGRTEYRRRRHTMDKDSGAA 69
 DB 14 SQTEADLALRPPLPLGTAGQ-----PRAGPP-----RRARFSK----- 49
 QY 70 ATTTEHFFRRSVCDNATLLEPLGLSLPQSPAPAVQAPPEPHREETATATA 129
 DB 50 AEP RPSSSRISRRSSVDLGLSSWSLPASP--APDPDP--PDSAGPBAR----- 96
 QY 130 TSGVAQPPAAPAGGAVAGPAPSTVPSTSDRVSQPSLVGSKPEPPPARSGSGGS 189
 DB 97 -----SPPSSKPPPGTWTGAP--VKAEDSARPELPDSAVGPSREP----- 139
 QY 190 AAEPOE---ERSQODDIELELTKAVGMSNGRFLFDIEIGSFYVKGGLDTEITYE 246
 DB 140 LRVPEAVALLERREGEKEKEDMETQAVATSPDRYLKFDIEIGSFYVKGGLDTEITYE 199
 QY 247 VAMCEIQDKLTSEQPRKEEEMKGLQHENVIVFYDSWSTVKGKCIYVLELMTS 306
 DB 200 VAMCEIQDKLTSEQPRKEEEMKGLQHENVIVFYDSWSTVKGKCIYVLELMTS 259
 QY 307 GLTKTYLRKFKYKIVLWSMCGQILKGLQHLTRRPPIIHRDLKNDNFIPTGPGSVKI 366
 DB 260 GLTKTYLRKFKYKIVLWSMCGQILKGLQHLTRRPPIIHRDLKNDNFIPTGPGSVKI 319

QY 367 GDLGIATLKRAASPAKSVIGTEPEFAPEYKEKDESDVVAFGMCMLEMAISEYYSBCQ 426
 DB 320 GDLGIATLKRAASPAKSVIGTEPEFAPEYKEKDESDVVAFGMCMLEMAISEYYSBCQ 379
 QY 427 NAAQIYRVVTSVGPVSPAPSPKVAIPVKEIIIEGCIROMDERYSIKDLINAPFOSETGVR 486
 DB 380 NAAQIYRVVTSVGPVSPAPSPKVAIPVKEIIIEGCIROMDERYSIKDLINAPFOSETGVR 439
 QY 487 VELAEEDGKELIAKLMIRIDIKKLGKYKONEAIEFSPFLERVPEDVAQEWESGVY 546
 DB 440 VELAEEDGKELIAKLMIRIDIKKLGKYKONEAIEFSPFLERVPEDVAQEWESGVY 498
 QY 547 CEGHKIMAKAIKRVSLIKRKEQROQLVREDEKKEQSESLKQVQSSAASQGTIKQL 606
 DB 499 CEAQYQVAVRVRERVAIQKREKLRARE-----LEAL 533
 QY 607 PSASTGIPTASTTSASVSTGV---EPEPEADQHQQLQYQOPST-SVLSDGTVNSGGSS 662
 DB 534 P-PEPGPPRATVPAPAPGPPSVFPEPEPEADQHQPLFRHASYSTTSDCTGYLSS 592
 QY 663 VFTESRVSSQQTWSYSGQHEQAHSTGTVPGHIPSITVAQSQPHGVYPPSSVAQSQSQGP 722
 DB 593 GFLDASDPALQP-----PGGVPSL--ASHLCLESA----- 622
 QY 723 SSSSLTVGVSSQPIQHPEQQQGIQCTAPPOQTVOYSLSQSTSSSEATTAQFVSPQAPQV 782
 DB 623 -----FALS----- 626
 QY 783 LPQVSAQSGQSTQGVAVAPAEVAVAAQPOATPQTLLASSVDSAHSDVAGSDGNENVP 842
 DB 627 IPRSGPSSDPSG-----DSVASDAASGLSDVGEEN-G 658
 QY 843 SSGRHGRTTKRYKRVSVRSRSHKTSRKLRLNYSNKGDRVVECOLFTHNKMTYFK 902
 DB 659 QMRPPGRNLR-----RPRSR-----LRVTSVSDQNDRVVEQLQTHNSKMTFR 704
 QY 903 FDLDDGNPEIATIMANDITLIERESFVDQVREITEKADENISGVSVEPBGQLES 962
 DB 705 FDLDDGSPFEEIAAMAYNEFILPSEBDFLRIREIILQVETELIKRDTGMMAEDTLSP 764
 QY 963 LQCKDYGFGSGQKLGEFPKQPIPASSMPQIGIPITSLQVYVHSAQRRTIVSPVPSRL 1022
 DB 765 QF-----EPAPLPALVPLVDPSNBEIQ-----SSTS 791
 QY 1023 RESKVPSEITDTVAASTAQSPGMNLHSAASLSLQCAFSELRAQWTEGPTAPPFVSH 1082
 DB 792 LEHRSWTAFSTSSSPOTPLSPG-NPFSPTPIPSGPIF-----PITSPP--CH 837
 QY 1083 TGPFTFVVPFLSSISAGVPTTAATAAPVATSSPPNDISTSVIQSEVYVTEBGIAGVAT 1142
 DB 838 PSEFS--PSPFISQVSNPSPHPTSSPLPSSSTPR----- 871
 QY 1143 SNGVVTSGGLPIPVESSEPVLSVSVSITIPAVVSISSPSSQVPTSEIYVSSATLY 1202
 DB 872 -----FPV-PLSQCP-----MSSLPTTSPPTSP-CSIQLSSPFR- 906
 QY 1203 PSVTYSATASAGGTATGPKRPAPVAVSOAAGSTTVGATILTSSTTGFPTASLOLSIQ 1262
 DB 907 -----PPC-----PSTSEPTTAAPLPS 925
 QY 1263 LSSSTSTPLAETVAVSAHSLDKTSHSSTTGAFSLISADSSSSSPGAVSSYISQPGGLH 1322
 DB 926 LSS-----AFSLA----- 933
 QY 1323 PLVPSVIASTPLPQAAGPTSTPLLQVPSIPPLVQVAVNPAVQQLIHSGQPALIP 1382
 DB 934 -----VMTYAAQSLL--SPSGGLIS 950
 QY 1383 NOPHTHCEVSDTQPKAGIDDIKTLERKLSLPSSEHSSGAQHASVSLFSTSVIESTV 1442
 DB 951 QSP----- 953

QY 1443 TPGIPTTAAPSCLTSTSTSTCLPTNLPLGTVALPVTPTGVSTVSTTSGVKEG 1502
DB 954 -----PA-----PPAPLP-----SLPLPPVAFGQESP-SPTIAVE-- 985
QY 1503 TABSKPPLTKAPVLPVGTETPAGTLESEQLPPFPFGSLVQSQPLEDLAQLRTLSPEK 1562
DB 986 -SEASPP-----PARLPGE----- 999
QY 1563 ITVTSVAVGVSMAPTAITEAGTQPOKGVSYQVEGVLATSSGAGVFKMGPFQVSVAADG 1622
DB 1000 ----ALAIAPIS-----EEGRPOL-----VGRFVYT----- 1020
QY 1623 AQEGKXKEDAKSVAFESSTSSSVLSSSPSTLVKEPENGITTPGISDVPESAHKT 1682
DB 1021 ----SSKEPAPPLPLQPT--SPTLSSG-----PKPSTPQLTSSSDTEJDA-- 1060
QY 1683 TASEAKSDTGOPTPKVGRFOVTTANKVGRFSVSKTDKTTDTYKEGVPASPPMDLEQAV 1742
DB 1061 ----GGGP-----ETRELAESDR----- 1075
QY 1743 LPATIVPKKEKPELSEPSHLNGSPSDPEAAFLSRVDVDGSGSPH---SPHOLSSKSLPSON 1799
DB 1076 ----AABGIGAGVBEEDGDK-EPYVGSGSPQLSHSPSEVMWN 1112
QY 1800 LSGSLNSFNSSWYSSDNESDIEDDLKELRLRLDKHLKEIQDLSRQKHETIESLYTKL 1859
DB 1113 VSYSS-----SLCISSESESGEDFEFMAELQSLRQHLSSEVETLQTLQKKEIEDLYSRL 1167
QY 1860 GKTPPAVILPPAPLGGRRRPTKSGS-KSSRSSSTIGKNSPOLSG-----NLSGGSAS 1913
DB 1168 GKPPPGIYAPVAPMALLSRQR--LSKGSFPTSRRNSLQSEPPGPGIMRRNSLSGSGTGS 1225

RESULT 10

Q810H4 PRELIMINARY; PRT; 1222 AA.
AC Q810H4;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE MNK4 Ser/Thr Kinase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MKY/NHsd; TISSUE=Kidney;
RA Toka H.R., Kahle K.T., Wilson F.H., Nelson-Williams C., Lifton R.P.;
RT "role of MNK4 and MNK1 in hypertensive rat strains."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY192567; AAC8858.1;
DR GO: GO:000524; F:ATP binding; IEA.
DR GO: GO:0004674; F:Protein serine/threonine kinase activity; IEA.
DR GO: GO:0004713; F:Protein-tyrosine kinase activity; IEA.
DR GO: GO:0006468; F:Protein amino acid phosphorylation; IEA.
DR InterPro: IPR000719; Prot. kinase.
DR InterPro: IPR002290; Ser_Thr_kinase.
DR InterPro: IPR008271; Ser_Thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot. kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TYRc; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50108; PROTEIN_KINASE_ST; 1.
KM Kinase.
SQ SEQUENCE 1222 AA; 132832 MW; E7E53DAC9D1FA446 CRC64;

Query Match 19.3%; Score 2083; DB 11; Length 1222;
Best Local Similarity 32.0%; Pred. No. 1.2e-92;
Matches 598; Conservative 167; Mismatches 356; Indels 750; Gaps 53;

QY 103 POPSTPAVPOSAPPEPHREETVTAATATNSQVAAQPPAAAP-----GEOAVAGPAP 153
DB 24 PPSPTSMGPTRLGPPPRVRFRFSGKABRPSSRRSSVDLGLSSWSQPSALLPBP 83
QY 154 STVPSTSKDRVQSPSLVGSKEP-----PPASGSGGSAKXP-- 193
DB 84 PDPSPSAGPWR--SPSS--NSKEHPGWTGAAPVKAADSACELTVSSGGPQSRPPRV 139
QY 194 ----GEERSQOQDDIEETKAVGMSNDGRFLKFDIEIGRSFKTYVKGLDTEETVEAV 249
DB 140 PDAARERRRQREKEDTETQAVATSPDGRVYKFDIEIGRSFKTYVKGLDTEETVEAV 199
QY 250 CELQDKLTYSRQRFKEAEMLKGLQHPNIYFIDSWESTYKCKKCIYLVTELMTSGTL 309
DB 200 CELQTKLSRAERQRESEVEMKGLQHPNIYFIDSWESTYKCKKCIYLVTELMTSGTL 259
QY 310 KTYLRKFKMYKLVLSKRCQLKGLQELHTTPPIIHDLDKDNFIETGPGSVKIGL 369
DB 260 KTYLRKFKMYKLVLSKRCQLKGLQELHTTPPIIHDLDKDNFIETGPGSVKIGL 319
QY 370 GLATLRASFAKSVIGTPPFMAPEMYEKEYDSVDVYAFGMCLMATSEYPSSECONAA 429
DB 320 GLATLRASFAKSVIGTPPFMAPEMYEKEYDSVDVYAFGMCLMATSEYPSSECONAA 379
QY 430 QIYRRTSGVKKPASFDKVALPEVKEITBECIFONKDERYSIIDLNAHAFQESTGYRVL 489
DB 380 QIYRRTSGVKKPASFDKVALPEVKEITBECIFONKDERYSIIDLNAHAFQESTGYRVL 439
QY 490 AERDDEKTAIKMLIEDIKKLKGKYKDNMAIEPSFDEBRDPEPEVADEMVSQVYCG 549
DB 440 AERDDEKTAIKMLIEDIKKLKGKYKDNMAIEPSFDEBRDPEPEVADEMVSQVYCG 498
QY 550 DKTKMAIKDRVSLIKRRRQROLVREOEKKQKQESSLSKQVESSASQTSIKOLPSA 609
DB 499 DYQVAVARVBERAAIQRKREKLRARE-----LEVLPPD 533
QY 610 STGIP-TASTSASVST-QVEPEPADQHOQLQYQPSI-SYLSGTYDSSGGSVYTFE 666
DB 534 SGPPATVAMTTPGPAPPEPEPEPADQHOQLQYQPSI-SYLSGTYDSSGGSVYTFE 533
QY 667 SRVSSQGYTSYSGQHQAISTGTVPGHIPSTVQAQSQPHGVPPSSVQAQSQGQSSSS 726
DB 594 ASDPALQP-----PGMPS--SPAEPLCLPSG-- 619
QY 727 LTGVSSSPPIQHQOQGGIQQTPAPQOQTQVYSLQSTSSSEATTAQVSPQAPQVLPQV 786
DB 620 ----FALS-----IPRS 627
QY 787 SAGKSTQGVSVQAPAEPAVAPQATQPTTLASSVDSASHDVASGMSDGENEVSSGR 846
DB 628 GPGSDSPG-----DSYASDAASGLSD-----M 650
QY 847 HEGRTKRRHRSKRSRSHREKTSRKRLTLYNSKNGDRVEQCLTHNRKMYTFKFDLD 966
DB 651 GEGQMKKPKVKTLLRR-----PKRLRTSVSDSDRVEQCLTHNSKMYTFKFDLD 704
QY 907 GDNPEELATIMVNDPILAIERESFYDQVEIIEKADEMLSEDVSVEPEGDQGLIESLOGK 966
DB 705 GDSPEELAAAMVYNEFILPERDGLSRIRREIIQVETLLKRD----- 747
QY 967 DDYGSQSQKLGEFKQPIPASSMPOQIGTISLTQVHASAGRRFLVSVPESRLRESK 1026
DB 748 ----AG----- 749
QY 1027 VFPSEITDYAATASTAGSPGNLSSASLSLQAFSELRRAQTEGENTAPPMFNSHTGPT 1086
DB 750 --PSEATEDALSPQEEPAAPALPGPSDALQSSISPEQSS----- 788
QY 1087 FVVPPEPLSSIAGVPTTAATAVAPVATSSPNDISTSVIGSEVTPTEBGIAVANSTGV 1146
DB 789 -----MAAFSTP-----SSP-----GPIPSPT 807
QY 1147 VTSGGLPIPEVSSPVLSSVSSITTPAVVISTTSBLOVPTSTSEIVVSTALPFSVT 1206

```

Db 808 PFSFG--TPVFPCPI-----FRTSSGCHP-----YFSSQ 836
QY 1207 VSASASAGSGSTAPFGKPPPAVVSQAAGSTTVATLTSTTSPFSLASQLSLSS 1266
Db 837 VS-----SNPCQAPSSLLPSSSGAQV-----PFS----- 863
QY 1267 TSTPTLAETVVSASHSLDKTSHSSTTGLAFSLAPSSSSPGAGVSYISQPGHLPLVI 1326
Db 864 ----- 863
QY 1327 PSYASTPPIILFOAAGPTSTPLLPQVPSIPPIVQPVANVAVQTLHSQOPALLPNOQH 1386
Db 864 PSL-----PTSSP-LPFSPSYP-----QVPLH-----PALTP----- 890
QY 1387 THCEVSDSDTPKAPAGLDDIKTLEKRLSEHSHSSGAGHVASLSTSLVIESVTPTGI 1446
Db 891 --CPS----- 897
QY 1447 PTTAVAPSKLLTSTSTCLPPTPLPLGTVAPVTVVPGQVSTPSTTSGVKGTAAPS 1506
Db 898 PSTAABLILASAFS-----LAVMTVA-----QSLISP-----SPGLISQ 933
QY 1507 KPPLTKAPVLPVGTPLPAGTLPSEQLPFPFGPSLTGSOQPLEDIAQLRRTLSPEXITVT 1566
Db 934 SPP-----APPGPLPSMPL-----PLACDQE----- 955
QY 1567 SAAGVMAAPTAIT--EAGTQPKGVSYQKEPVLATSSGAGVFKKGRQVSVADGAQ 1624
Db 956 -----SLSAQIATEENEARNPQO-----PLLDARLAPI----- 985
QY 1625 KEGKNSKEDAKSVHESSTSESSSVLSSSPESTLVKPEPNTITPGISDPVESAHTTA 1684
Db 986 ----- 985
QY 1685 SEAKSTGQPTKYGRFOVTTANKVGRFSVSKTEKIDTDTKKEGPVAPSPFMDLEQAVLP 1744
Db 986 -----SEEGKQVLGRFQVTS-----KE--PAEPPLQASPTLSR 1019
QY 1745 AVTPKKEKPEL-----SEPSHLNGP-----SDPEAFSLRDVDGSGSHSPQLS 1791
Db 1020 SL--KLPTEPQLTSESDTESSAAGPETREALAESRAAGLGVAILDEKDEKEPIQG 1077
QY 1792 SKSLPSQNTLSQSLNSFNSSSYMS--DNESDIEDDKLKLRLRDKHLKELIODLSROKH 1850
Db 1078 SPSLLQSPVVMNYSYSLCLSSSESGDEBEFAHLQNLROKHLSEVALQTLQKK 1137
QY 1851 EIESLYTKLGKVPVAVIIPPAAPLSGRRRPTYSKGS--KSSRSSLSGNKSPOLSG----- 1904
Db 1138 EIEDLYSRGKOPPPGIVAPAAMLSCORR--LSKGSFPTSRNSL--QESDLPFGCIMG 1193
QY 1905 --NLGOSAS 1913
Db 1194 RNLSSSSTGS 1204

```

RESULT 11

```

O811RS ID O811RS PRELIMINARY; PRT; 1210 AA.
AC O811RS;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN PRKNMKA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WKY; TISSUE=Kidney;
RA Zimdahl H., Monti J., Hubner N.;

```

```

RT "Sequence analysis of WKNA in spontaneously hypertensive rats.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY187039; AAO18238.1; -.
DR PIR; A40135; A40135.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein tyrosine kinase activity; IEA.
DR GO; GO:0004648; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein.
SQ
SEQUENCE 1210 AA; 131499 MW; 1C98471497DEC65B CRC64;

```

```

Query Match 19.3%; Score 2082; DB 11; Length 1210;
Best Local Similarity 31.8%; Pred. No. 1.3e-93;
Matches 593; Conservative 168; Mismatches 364; Indels 742; Gaps 50;

```

```

QY 103 PPSIAPAVQSPAPPPHRETYTATNSQAQPPAAAP-----GEQAVAGAP 153
Db 12 PPSITSMGPTLGLPPRRRRPFGKAPPRPSRSRSDVDLGLLSSWQPSALLPEP 71
QY 154 STPSSSTSKDPRVQSOLVSKKEP-----PPASGSGGSGAKEP-- 193
Db 72 PDPDDAGPKR--SPSS--NSKEHPGWTGAAPVYAAADACELTVSSGPGSGREPRV 127
QY 194 ----QEERSQQDDIELETKAVGMSNGRFLKFDIEIGRSFRTYKGLDTEETVEAV 249
Db 128 PDAARERREGEKEKEDTEQAVATSPDGRYLFKFDIEIGRSFRTYKGLDTEETVEAV 187
QY 250 CELQDRKLTKSEQRKEPEMKLGLQHPIVAFYTSWSEYVKGKCIYVLEMLTSGTL 309
Db 188 CELQARKLSRAEQRSEEMKLGQHPNIVAFYDSWKSVALRGVCIYVLEMLTSGTL 247
QY 310 KYLKRKFYWKIKVLSWCRQILKGLQPLHTRPPTIHHDKDNTFITGPGSVYIGLU 369
Db 248 KYLKRFRFKRRVLDQMSRQILKGLHFLHSRVPPLIHHDKDNTFITGPGSVYIGLU 307
QY 370 GLATLRASFASVIGTPEFMAPEMYEKYDESVDVYAFGMCMLENATSEYPSYSECQNA 429
Db 308 GLATLRASFASVIGTPEFMAPEMYEKYDEAVDYAFGMCMLENATSEYPSYSECQNA 367
QY 430 QIYRYTSGVKASPPKVAIPPEKELIEGCIQNKREKRSIKQLNHAFFQETGYRVEL 489
Db 368 QIYRYTSGTKPNSFYKVMPEVKELIEGCIQNDKNERFTIIDLTHAFREERGVHVEL 427
QY 490 AEDDGEKIAIKMLRIEDIKLKGKYKQNEAIEFSFDELDVPEPDAQOMVEGVCYCE 549
Db 428 AEDDGEKPGIKMLRMEDARR--GGRPRNQAIIEFLFQGRDAEVEADAMVALGLVCE 486
QY 487 DYQPARAVRERVAAIQREKELIKARE-----LEVLPPD 521
QY 610 STGIP--TASTTASVST--QVEPEEPADHQLOQVQPSI--SVLSNGTVDSGQSSVFTE 666
Db 522 SGPPPATVSMTPGPPSAPFPEPEEPADHQSLFEHHAYSSTSDSCETDGYLSSGFLD 581
QY 667 SRVSSQQTIVSYGSOHEQASHSTGVPGHLPSTVQAOQSPGVVPPSSVAGQSQGQSPSSS 726
Db 582 ASDPMLP-----FGGMS--SAERHLCLPSC----- 607
QY 727 LTGVSSQPTIQHPQQQGGIQQTAPPOQTVQVLSQTSSTSSAATTAQVSPQAPQVLPQV 786
Db 608 -----FALS-----IPRS 615

```

QY 787 SAGKOSTQGVSOVAPAEPAVAVAPQATQPTTLASSVDASHASGMSGDNENVPSSSGR 846
 Db 616 GPGSDPSFG-----D5YASDPAASGLSD-----M 638
 QY 847 HEGRTTKRHVRKSVRSRSHKTSRPKRLINYSNKGDRVVEGQLTHNRKMTVEKFD 906
 Db 639 GEGGQKQKQVKTLLRR-----PSSRLRVTSDSDSDRVEQLQTHNSKMTVEKFD 692
 QY 907 GDNPEIATIMVNDFTLAIERESFVDQVREIEKADDEMISEDVSVPEQDQGLSLQCK 966
 Db 693 GDSPEIILAMVNEFLPESRDEGLSRIRREIQVETLLKR----- 735
 QY 967 DDYFGSGQKLEGEFKQPIPPASSMPOQIGIPSSILQVHNSAGRPVSPVESRLRESK 1026
 Db 736 -----AG----- 737
 QY 1027 VPPSEITDVAATASTASPGMNLSHSASLSLQQAPELRAQGTGEPNTAPNPSHTGPT 1086
 Db 738 --PSEATEDALISQGEAAMPALPPSDELORSISPEORS----- 776
 QY 1087 FVVPVPELSSIAVPTTAATAAVPAPATSSPPNDISTSVIQSEVTPTEEGIAGVATSGV 1146
 Db 777 -----WAARSTSR-----SSP-----GTPLSPGT 795
 QY 1147 VTSGILPPIPVSESPVLSSVSSITTPAVVSISSPSLQVPTSTSEIIVSSITALYPSVT 1206
 Db 796 PESPFG--TPVPVPCPI-----PFTSPSCHP-----YPFQ 824
 QY 1207 VSAITSAAGSSTTPQKPPAVVSOQAAGSTYGALITVSTTTSPPTASQSLQSS 1266
 Db 825 VS-----SNPCQAPSSILPSSSGASQV-----PPPS----- 851
 QY 1267 TSTPTLAETVWVAHSLDKTSHSSTGLAFSLAPSSSSPGAGVSVISQPGCLHPLVI 1326
 Db 852 ----- 851
 QY 1327 PSYIATSTPLPQAGTSTPLPQVPSIPPLVQPVANVAVOQTLIHSPQAPALLPNOQH 1386
 Db 852 PSL-----PTSSP-LPFPSPSY-----QVPLH-----PASLPT-- 878
 QY 1387 THPEVDSDTQAPAPGIDIKLEKRLRFSHSSSGAOMASVLSFTSLVHSTYTPGI 1446
 Db 879 --CPS-----PPPL 885
 QY 1447 PTTAVAPSKLITSTSTCLPPTNLPLGTVALPVTVPVTPQOVSTPVSTTSGVKPCTAPS 1506
 Db 886 PSTTAPPLSLASAFS-----LAWMTA---QSLTSP-----SPGLSQ 921
 QY 1507 KPPLTKAPVLPGVTELPAGTLPSEQLPPPPPSLITQSQQPLBDLDAQLRRTLSPEXITVT 1566
 Db 922 SPF-----APGPLPSMPL-----PLASCQGE----- 943
 QY 1567 SANGVPSMAPTAIT--EAGTQPKGVSOYKEBPVLTSSGAGVFKMGRFQVSVAADGAQ 1624
 Db 944 -----SLSAQTAEITENASRNPAQ-----PLGDALAPI----- 973
 QY 1625 KEGKNSKEDAKSVHFESSTSSSVLSSSPSTLVKPEPNTGITPGISSDPESAKHTTA 1684
 Db 974 ----- 973
 QY 1685 SEAKSDTQOPTKVGROVTTANKVGRFVSXTEDKITDTKKGPAVAPPFMDLEQAVLP 1744
 Db 974 -----SSEKQQLVGRFOVTS-----KE--PAEPPLQASPTLSR 1007
 QY 1745 AVLPKKEKEL-----SEPSHLNGP-----SSPPEAFLSRVDVDSGSHSPQLS 1791
 Db 1008 SL--KLPTQLTSESDTDESAAGPSTRALAESRAAGLGAVIDEEDKEKEQIG 1065
 QY 1792 SKSLPQONLSQSLSNFNSGYMS--DNESDIEDDKLTLRLRDHKLKEIIDLQSRKH 1850
 Db 1066 SSSILSQPSVWMMNYSSLSLCSSESSESGEDEFPAHLQNLQKHLSEVALQTLQCK 1125
 QY 1851 EIESLYTKLGKVPDAVILPPADPLSGRRRRPTKSGKSSRSSSLGNKSP---QLSGNL 1906

Db 1126 EIEDLYSLGKQPPGIVAPAAMLSCRRRLKSGSFPTSRNLSLQSDLPFGIMKNSL 1185
 QY 1907 SCSAAS 1913
 Db 1186 SCSSTGS 1192
 RESULT 12
 080UE6
 ID 080UE6 PRELIMINARY; PRT: 1222 AA.
 AC 080UE6,
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE WNK4.
 GN PRKNK4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RA MEDLINE=22426867; PubMed=12515852;
 RA Wilson F.H., Kahle K.T., Sabath E., Lalicci M., Rapson A.K.,
 RA Hoover R.S., Hebert S.C., Gamba G., Lifton R.P.,
 RT "Molecular pathogenesis of inherited hypertension with hyperkalemia:
 RT The Na-Cl cotransporter is inhibited by wild-type but not mutant
 RT WNK4".
 RT Proc. Natl. Acad. Sci. U.S.A. 100:680-684(2003).
 DR EMBL: AY187027; MAF21955.1; -
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:004674; F:protein serine/threonine kinase activity; IEA.
 DR GO: GO:004713; F:protein-tyrosine kinase activity; IEA.
 DR GO: GO:006468; F:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR000719; Prot. Kinase.
 DR InterPro: IPR002290; Ser. thr. kinase.
 DR InterPro: IPR008271; Ser. thr. pkin. AS.
 DR InterPro: IPR001245; Tyr. kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot. Kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TyrK; 1.
 DR PROSITE: PS00101; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
 SQ SEQUENCE 1222 AA; 132409 MW; BE0A193E7D79DCBE CRC64;
 Query Match 19.1%; Score 2068; DB 11; Length 1222;
 Best Local Similarity 31.8%; Pred. No. 6, 6e-93;
 Matches 601; Conservative 167; Mismatches 352; Indels 772; Gaps 55;
 QY 97 GRLS-----LPPSIPAAVPGSAPPEPHREIVTATASQVAAQPPAAAP----- 143
 Db 10 GVPMSQTEADLALRSPALTSPTPLGPPRRVRRFSKABPRSSRSRSDVLGL 69
 QY 144 -----GEOAVAGPAPSTVPSSTSKDRPVSOPLSVGSKKEPP-----PAR 182
 Db 70 LSSMSQPAALLPFPDPDPSAGPTR--SPPS--SKPEPBGTMGAAPVXAVDSACPL 124
 QY 183 SSGGGG-SAKEP-----QEERSQQODDIIELETYAVGMSNDGRELFKPDIEIGRSFKT 234
 Db 125 TSSSGGPGSREPRVPDAARERRRRQKEKEDTETQAVATSPDGRYLKPDIEIGRSFKT 184
 QY 235 VYKGLDTETTVANWELQDRKLTSEORQFKAEAMLGLOHPNIVREFYDSKSVLRQ 294
 Db 185 VYKGLDTETTVANWELQDRKLTSEORQFKAEAMLGLOHPNIVREFYDSKSVLRQ 294
 QY 295 KCIIVLTETMTSGTLTKYTKRKVKMKIYVLRSGCROILKGLQPLHTPTPIIHRDLKCN 354
 Db 245 VCIVAVTETMTSGTLTKYTKRKVKMKIYVLRSGCROILKGLQPLHTPTPIIHRDLKCN 304
 QY 355 IFITGFTGSKIGDGLATLTKASPAKSVIIGPEFMAPEMYEKKYDESVDYVAFGCMLE 414

```

Db      305 VFTTSGSGSVKIDGLIATLKRAFAKSVIGTPEPFMAPEMYEEXDEADVAFGMCMLE 364
      415 MATSEPYSSCQNAQAQYRRTVSGVKAAPDKXALPPEVKIIEBCTIRONDERXSTODL 474
      365 MATSEPYSSCQNAQAQYRRTVSGVKAAPDKXALPPEVKIIEBCTIRONDERXSTODL 424
      475 NHAFFORETGVRLAEEDDEGEKIALKMLRIEDIKLKGKYNONEAIEPSFDLDPVE 534
      425 AAHFFRERGVHLEAEDDEGEKGLKMLRMEDAR--GGRPDNQAIEFLFQGRDAE 483
      535 DVAQEMVSESYVCEBGDKHTAKAIKORVSLIKRRREGQIVREOEKKQESSLKQVE 594
      484 EVAQEMVALLVCEADYQVAVARVERVAIQRREKLRKARE----- 526
      595 QSSASQGTGIKQLSASTGITASTTASVSTQV---EPEEPADQHQLOLOQOQPSI--SVL 650
      527 -----LEVLPPDS--GPPPAVLSLAPGPPSAFPPPEPEEADQHQSLFRHASYSTT 577
      651 SDGTVDGGQSSSVFTESRVSQOTVSYGQHEQAHSTGVGHIPSTVQAQSPHGYPP 710
      578 SDCEETGYLSSSGFLASDPALQP-----PGLPSS----- 608
      711 SSVAQGGQSGQSPSSSLITGVSSSQPIQHPOQOQGIQGTAPPOQTVOYSLSTSSBATT 770
      609 -----PAESHLCPLPSG-----FALS----- 623
      771 AQVSPQAPQVLPQVYASAGKQSTQGVAPAPPAVNAQAPQATQPTTLASVDSANSDVA 830
      624 -----IPRSGPSPDFSPG-----DSYASDA 644
      831 SGMSDGENENVPSSSGHREKRTTKRHKYSVRSRHEKTSRPKILRLYNKNGDRVCEQ 890
      645 SGLSD-----MGEQGQMKNPVKTLRR-----PRSLRLRTVSYSDQDRVCEQ 688
      891 LEHNRKMTFKEDLDGDNPEEATIMWDDFLAIBRESFVDQVREIEKADMLSEVY 950
      689 LQTHNKRMTFKEDLDGDNPEEATIMWDDFLAIBRESFVDQVREIEKADMLSEVY 747
      951 SVPEBGDQGLSELDGKDDYGFSGSQKLEGEFKQPIAASSHPQOIGIPTSLTLQVHSAQR 1010
      748 ----- 749
      1011 RPIVSPVESRLRESKVFPEITDTVAASTAGSPGMNLSHSAASLSLOQAFSELRAQMT 1070
      750 -----PPE-----AABALSP-----Q 761
      1071 EGPNTAPPNFSHTGTPFPVPPFLSIAGVPTTAATAAPVATSPSPNDISTSVIOSEVT 1130
      762 EBPALP-----ALPQPPMAEPQRSIS----- 783
      1131 VPTEBGIAGVATSTGVVTSGLPIPPVSESFVLSSVSSITIPAVVSIETPSLQVPTS 1190
      784 -PEQRMAAFSTSP--SSEGTPLSP----- 805
      1191 TSEIVVSTALVPSVTSATASAGSSTATPGKPPAVVSOQAAGSTTGATLTLSVSTTT 1250
      806 -----CAPSPGPP----- 815
      1251 SEPSTASQSLQSLSSSTPTTAAETVAVVSAHSLDKTSHSSTGLAFSLAPSSSSSPGAG 1310
      816 VEPSP-----IPPIPSPCYCPGPFQ 836
      1311 VSS--YISQGGHPLVIPSVLASTPILPQAAGPTSTPLLPQVPSLP--PIVQPVANVA 1366
      837 VSSNPFPQAPSSLPPL--SSASQVPLPSSSLPISAP--LPSSPSYPOPL----- 883
      1367 VQQTLLHSQPOPALLPNOPHTHCEVDSDTQPKAPGIDIKLIEKRLSLFSEHSSGQAQ 1426
      884 -----SPTSLSL-----VCPSS----- 893
      1427 HASVSLSTSLVLESTVTPGIPPTTAAVPSKLLNSTSTCLPEPTNLPLGYVALPVTVPVTPG 1486

```

```

Db      894 -----PPLSPSTTAAPLSLASAFS-----LAVMTVA---QSLSP- 926
      1487 QVSTPSTTTSGVKGPTAASKPLTKAPVLVGTETLPACTLPBEQPLPFPFGPSLTOSQOP 1546
      927 -----SPGLSQSP-----APPGPLPSL----- 945
      1547 LEDLDQLRRTTSPKXITVTSVAVPSMAAPTAITEAGTPQKGVSGVKGVPVLTSSGA 1606
      946 -----PLSLASCD----- 953
      1607 GYFKMRFPQVVAADAQKQKNSBDKSVHESSTSESSVLSSSPESTLVKPEPNGI 1666
      954 -----GESLSAQIAETENE-----ASNPAPPL----- 977
      1667 TIRGISDVPESHAKHTASAK-----SDTGQPTKVGAFQVTTANAKVGRSVKTEKDXT 1722
      978 -----GDARLAPISEBKPOLVGRFOVTS----- 1002
      1723 DTKKEGPAVSPPPMDLEQAV-----LPAVLPKKEKELSPSHLNGP-----SSDPEA 1770
      1003 ---KE--PAEPPLQAPSPILSRSLKLPSP--PLTSESDTSDSAAGPETEALABSDRAA 1056
      1771 AFLSRVDGSGSPHSPHOLSSKSLPSQNLQSLSNFSNYSYS--DNESDIEDDLKTE 1829
      1057 EGLGVAVDEKDEKGPPLIGSSSPILSHSPVMMNYSYSLCSSESSSGEDDEEFMAE 1116
      1830 LRRLRDKHLKEIODLOSQKHETESLYTKGKPPRAVILPPAPLGRRRRPFKSGS--K 1888
      1117 LQMLRQKHLSEVALQTLQKKEILEDYLSRIGKPPGIVAPAMLSRCORR--LSKGSFP 1174
      1889 SSRSSSLGNKSPQLSG-----NLGQSAAS 1913
      1175 TSTRNLSL--QRSDLPQPGIMRRNLSLSSSSTGS 1204

Db      1175 TSTRNLSL--QRSDLPQPGIMRRNLSLSSSSTGS 1204

RESULT 13
Q8N673
ID Q8N673 PRELIMINARY; PRT; 393 AA.
AC Q8N673;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RU Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021121; AAH21121.1; -
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Hypothetical protein; ATP-binding; Transferase.
SQ
SEQUENCE 393 AA; 42434 MW; 89555C4EF15BFD3E CRC64;

Query Match 18.3%; Score 1974; DB 4; Length 393;
Best Local Similarity 100.0%; Pred. No. 5.5e-89;

```


Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGAIAEKOSSTGSLFLSPAPAPKNGSSDSVGEKIGAAADAVTGRTEYRRRRRT 60

Db 1 MSGGAIAEKOSSTGSLFLSPAPAPKNGSSDSVGEKIGAAADAVTGRTEYRRRRRT 60

QY 61 MDKDSGAAATTTTTHRRFRFRSVICDSNATLELPLSLPQPSIPAAPVQAPDEPH 120

Db 61 MDKDSGAAATTTTTHRRFRFRSVICDSNATLELPLSLPQPSIPAAPVQAPDEPH 120

QY 121 REETVATATSOVAQOPPAAPAAAGVAVGAPASTVPSSTSDCRPSQPSLVGSKPEPP 180

Db 121 REETVATATSOVAQOPPAAPAAAGVAVGAPASTVPSSTSDCRPSQPSLVGSKPEPP 180

QY 181 AASGSGGSAKEPQERSSQODDIELETKAVGMSDGFLEKFDIEIGSFYVYKGD 240

Db 181 AASGSGGSAKEPQERSSQODDIELETKAVGMSDGFLEKFDIEIGSFYVYKGD 240

QY 241 TETTVAVANCELODRKLTSEKQFEAEMLKGLQHPNIVRFYDSWSTVKGKCIIV 300

Db 241 TETTVAVANCELODRKLTSEKQFEAEMLKGLQHPNIVRFYDSWSTVKGKCIIV 300

QY 301 TELMTSGTLKTYLKRKKVWKIKYLRSCQILKGLQFLHTRPPIIHRDLKCDNFIITGP 360

Db 301 TELMTSGTLKTYLKRKKVWKIKYLRSCQILKGLQFLHTRPPIIHRDLKCDNFIITGP 360

QY 361 TGSVKIGDGLATLKRSPAFKSVYIG 385

Db 361 TGSVKIGDGLATLKRSPAFKSVYIG 385

RESULT 14

Q80XN2 PRELIMINARY; PRT; 1048 AA.

AC Q80XN2; 24, Created

DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE 2010002J1IRIK protein (Fragment).

GN GN 2010002J1IRIK

OS Mus musculus (mouse).

OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

NCBI_TaxID=10090;

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N; TISSUE=Kidney;

RA MEDLINE=22388257; PubMed=12477932;

RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marinina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stappleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,

RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimm J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,

RA Jones S.J., Maitra M.A.,

RT "Generation and initial analysis of more than 15,000 full-length human

RL and mouse cDNA sequences."

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N; TISSUE=Kidney;

RA Klausner R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC043677; AAH43677.1; -.

DR GO: GO:0005524; F:ATP binding; IEA.

DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.

DR GO: GO:0004686; F:protein amino acid phosphorylation; IEA.

DR InterPro: IPR000719; Prot Kinase.

DR InterPro: IPR002290; Ser thr kinase.

DR InterPro: IPR008271; Ser thr pkinase.

DR InterPro: IPR001245; Tyr pkinase.

DR Pfam: PF00069; pkinase; 1.

DR ProDom: PD000001; Prot. kinase; 1.

DR SMART: SM00220; S_TKc; 1.

DR SMART: SM00219; Tyrc; 1.

DR PROSITE: PS00011; PROTEIN KINASE DOM; 1.

DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.

FT NON_TER

SQ SEQUENCE 1048 AA; 113942 MW; 45F8A8A6D18DA66 CRC64;

Query Match 17.7%; Score 1911.5; DB 11; Length 1048;

Best Local Similarity 32.1%; Pred. No. 2,5e-85;

Matches 551; Conservative 146; Mismatches 297; Indels 725; Gaps 48;

QY 228 GGSFRTYKGLDITTEVAVANCELODRKLTSEKQFEAEMLKGLQHPNIVRFYDSW 287

Db 4 GGSFRTYKGLDITTEVAVANCELODRKLTSEKQFEAEMLKGLQHPNIVRFYDSW 287

QY 288 ESTVKGKCIIVATLEMTSGTLKTYLKRKKVWKIKYLRSCQILKGLQFLHTRPPII 347

Db 64 KSVLKGQCVIVATLEMTSGTLKTYLKRKKVWKIKYLRSCQILKGLQFLHTRPPII 347

QY 348 RDLCNDNFIITGPTSVKIGDGLATLKRSPAFKSVYIGTPEFADMEYEKDESIVYA 407

Db 124 RDLCNDNFIITGPTSVKIGDGLATLKRSPAFKSVYIGTPEFADMEYEKDESIVYA 407

QY 408 FQKMLEMATSEYPSSECCNAQIYRRTSVGVKPSAFDVAIPEVKEITEGCRONKDER 467

Db 184 FQKMLEMATSEYPSSECCNAQIYRRTSVGVKPSAFDVAIPEVKEITEGCRONKDER 467

QY 468 YSIKDLNNAFPOEFGVAVELAEEDDGEKIAKLMIRIEDIKKLGKXKNDIAEFSPD 527

Db 244 FTIQLDLNNAFPOEFGVAVELAEEDDGEKIAKLMIRIEDIKKLGKXKNDIAEFSPD 527

QY 528 IERDVEDVADQVMSGVYCEGDHKTMAKIDRYSILKRKEQRLVREDEKKEKQES 587

Db 303 IGRDAEEVADQVMSGVYCEGDHKTMAKIDRYSILKRKEQRLVREDEKKEKQES 587

QY 588 SLKQVQSSASQGTGKQIPASSTGIPATSTTASAVSYV---EPPEPADQHOLOLYQ 644

Db 353 SLKQVQSSASQGTGKQIPASSTGIPATSTTASAVSYV---EPPEPADQHOLOLYQ 644

QY 645 PSI-SVLSPGTVDSGGSSVFPSRSVSSQGTYSQSHQAHSTGTPGHIPSTVQAQSQ 703

Db 397 PSI-SVLSPGTVDSGGSSVFPSRSVSSQGTYSQSHQAHSTGTPGHIPSTVQAQSQ 703

QY 704 PHGVYPPSSVAGQSGQSSSLTGVSSQPIQHPOQOQGIQQAAPQQTQVYSISQTS 763

Db 435 PHGVYPPSSVAGQSGQSSSLTGVSSQPIQHPOQOQGIQQAAPQQTQVYSISQTS 763

QY 764 TSEATTAQPVQOPAPQVLPVQVSAKQSTQGVQVAPAPVAVAPQATOPTTLASVSD 823

Db 450 TSEATTAQPVQOPAPQVLPVQVSAKQSTQGVQVAPAPVAVAPQATOPTTLASVSD 823

QY 824 SAHSDVAGSMDSGNGENRVPSSSRHEGRTTKRHYKSVGRSRHEKTSPEKRLILVNSKG 883

Db 464 SAHSDVAGSMDSGNGENRVPSSSRHEGRTTKRHYKSVGRSRHEKTSPEKRLILVNSKG 883

QY 884 DRVVECOLETHNRKVTFFKFDLDGDNPEEIIATIMVNDIFILARESFDVQVEIIEKAD 943

Db 508 DRVVECOLETHNRKVTFFKFDLDGDNPEEIIATIMVNDIFILARESFDVQVEIIEKAD 943

QY 944 EMLSEVSVPEBGDGLSLQKQDYGSQSLKEGFROPPIPASSMPOQIGLPTSLTQ 1003

Db 568 EMLSEVSVPEBGDGLSLQKQDYGSQSLKEGFROPPIPASSMPOQIGLPTSLTQ 1003

```

QY 1004 VHSAGRRFIVSPVPSRLRESKVPEITDVTVAASTAOSPGMNLSSHASLSLQQAASE 1063
Db 574 ---AG-----PPE-----AAEDALSP-----586
QY 1064 LRBAOMTEGENTAPNPNHSHTGPTFPVVPPLSSIAAGVPTTAATAAPVATSPENDISTS 1123
Db 587 ---QEEPAALP-----ALPDPRAAEQORIS--609
QY 1124 VIOSEVTPPEBEGIAVATSTGVVTSQGLPIPVSESPVLSVVSSITTPAVVISITSTP 1183
Db 610 ---PQORSWAFAFSTSP---SSPGTPLSP-----631
QY 1184 SLQVPTSTSRIVVSTALYPSVTVAATSAAGSTATAPOKPPAVVSOQAAGSTTVGATL 1243
Db 632 ---CAPSPSPGTP-----641
QY 1244 TSVSTTSPSTASQSLSIQSSSTPTLAEVTVVSAHSLDKTSHSSTTGIAFSAPSS 1303
Db 642 ---VFPCP-----IFPITSPEC 655
QY 1304 SSSPGAGVSS--YISQPGHPIVIVSVIASTPLIQAAGPTSTPLIPQVSTP--PLVO 1359
Db 656 YPCFESQVSSNPYPQAAPSSILPL---SSSASQVPLPSSSLPIEAP-LPFSPPSYQODPL--709
QY 1360 PVANVAVOQOTLIHSQOPALLENOPHTHCPEVDSPTQPKAPGIDDIKTLERLSLFE 1419
Db 710 ---SPTSLEP---VCPSS-----719
QY 1420 HSSSGAQAHSVLSLETSLVIESTVPGITPTTAVAPSKLITSTSTCLPPTNLPLGTVALPY 1479
Db 720 ---PFSLEPTTAAPLILSLASAFS-----LAWMTVA--746
QY 1480 TPVYTPGQVSTPVSTTSGVKPCTASKPPLTAPVLPVSTELPACTLSEQLPPEPPGS 1539
Db 747 QSLTSP---SPGLISQSP-----APPGPLPS 770
QY 1540 LTOSQOPLDLDAQLRRLTSPEKXITVTSAGVPMAPTAITEAGTQOPQGVQVKEGPV 1599
Db 771 L-----PLSLASCD-----779
QY 1600 LATSSGAVFKMGKRFQVSVAAADQAQKGNKSEDAKSVMHESSTSESVLSSSSPSTIV 1659
Db 780 ---OESLSAQTAETENE-----ASNPAQPLL 803
QY 1660 KPEPNGITPGISSOVPESAHKTASEAK---SDTQPTKVGKRFQVTTANKVGFVS 1715
Db 804 ---GDARLAPISEEGKPOLVGRFOVTS--828
QY 1716 KTEDKITDTKKEGPVAPSPFMDLEQAV---LPVAPIKKEKPELSEPSHLNGP-----1764
Db 829 ---KE--PAEPPLQAPASPTLSSILKLPSP-PLTSESDPTEDDAAGPETEREAL 875
QY 1765 -SSDPPAALPLSRVDVDDSGSPHSFOLSSKSLPSQNLSQLSNSFNSSYMS--DNESDIE 1822
Db 876 AASDRAAEGLGAVAVDDEKDEKPELIGSSSPILSHSPVMWYSYSLCISSESSSCE 935
QY 1823 DEDLKLRLRLRKHKEIODIOSROKHELESYTLGKVPVAVITLPPAALPSGRRRRT 1882
Db 936 DEEFMAELONLROKHLSEVALQTLQKKEIEDIYSRLGKOPPGIAPAAMLSCRQR--993
QY 1883 KSKGS-KSSRSSSLGNKSPOLSG-----NLSGQSAAS 1913
Db 994 LSKGSFPTSRNLSL--QRSDLPQGITMRKNSLSGSSSTGS 1030

```

```

DE AC2-059.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,
RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RT "Liver regeneration after PH.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY323328; AAP66260.1; -
SQ SEQUENCE 1552 AA; 173188 MW; 23AA592C8CDD088CA CRC64;

Query Match 17.5%; Score 1890.5; DB 11; Length 1552;
Best Local Similarity 42.7%; Pred. No. 4,6e-84;
Matches 452; Conservative 125; Mismatches 236; Indels 245; Gaps 28;

QY 103 PQSIPAAVQGAAPPEPHHEVTAATQVAAQPPAAAP-----GEQAVAGAP 153
Db 185 PPSPLTSMGPTRLGPPRRVRRPFGAEPRRSPRSRSDVDLGLSWSQASLLEP 244
QY 154 STVPSSTSKDRPVSOPLVSKKEP-----PPARSGSGGSAKEP--193
Db 245 PDPPDSAGPMP--SPSS--NSKEHPGVTGGAAPVAAADSACELTVSSGGPSRPPRV 300
QY 194 ---QERSQOQDDIELEETKAVGMSNGRFLKEDIEIGRGSFVYKGLDTEVAVM 249
Db 301 PDAAARERREDEKEDTQAVATSPDRYLKFDIEIGRGSFVYKGLDTEVAVM 360
QY 250 CELQORPKLKSROPKKEAEMLKGLQHNIYRFYMSWSTYKGGKCIYLYTEIMSSGL 309
Db 361 CELQTRKLSRARQRSEVEMLKGLQHPNIVRFYMSWSTYKGGKCIYLYTEIMSSGL 420
QY 310 KYLKKFKYMKIKVLRSMCRQILKGLQFLHTRPTPIIHRDLKDNFTIGTSPVXIGDL 369
Db 421 KYLRRFRMRKRVLRQWRQILRGLHFLSHVPLIHRDLKDNFTIGTSPVXIGDL 480
QY 370 GLATLKRAFPASVIGTPEFPAPEMYEKEYDSVDVYAFGMCKLEKATSEYPSSECONA 429
Db 481 GLATLKRAFPASVIGTPEFPAPEMYEKEYDAVDVYAFGMCKLEKATSEYPSSECONA 540
QY 430 QYRRTSGVKKASPEKVAIPVKEIIEGICRQNDERISINDLNHAFQESTGVREL 489
Db 541 QYRKVTSGTENSTFYKVMPEVKELIBGLIKTDKNERITQDLTHAFRBERGVHVEL 600
QY 490 AEEDGKIAIKMLRIEDIKLKQYKDNEMAIIEFSFDELPDVPEDVAQEMWESGYVCEG 549
Db 601 AEEDGKIKGLMLRMEDARR--GGRPRNDQAIIEFLFQGRDAAEVDAQEMVALGLVCEA 659
QY 550 DKTMAKAIKDNVSLIKRRRQOLVREBOEKKKQOESSLKQOEOSSASQTSIKOLPSA 609
Db 660 DYQPARAVAREVVAIQREKELIRARE-----LEVLPPD 694
QY 610 STGIP-TASTTASVT--QVEPEPEADQHQLOQYQOPI--SVLSPGTVDGSGSSVFPE 666
Db 695 SCPPPATVSMTGPPSAPFPEPEPEADQHQFLFRHASYSTTSCEFDGYLSSGFLD 754
QY 667 SVSSQGVTSYSGOHEAHSTGVPHLPTSTVQAOSQPHGVVPPSSVAAQSGQSQPSSSS 726
Db 755 AADPALQP-----PGWMS--SPABRP-LCLSVFALSIKSGGS--792
QY 727 LTGVSSQPIQHPOQOQIGQTAPQOQVYLSQTSISSEATTQAPVYQAPQVLPV 786
Db 793 ---DPSPGDMVYLVSVWQAT-----PQWQO 816
QY 787 SAGKSTGQGVAAPAEPAVAQAQATPTTLASSVDSHSDVYASGMSGNENWSSSGR 846
Db 817 ---ALSDMG-----822
QY 847 HEGRTTKHRYKSVRSRHEKTSRPKLRILNVSKGDEVVACOLETHNRKMTTEKFDLD 906

```

```

RESULT 15
Q7TPK6 PRELIMINARY; PRT; 1552 AA.
AC Q7TPK6:
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, last annotation update)

```


Blank

Sheet

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 25, 2004, 01:25:13 ; Search time 11801 Seconds

(without alignments)
7845.159 Million cell updates/sec

Title: US-10-010-720-14

Perfect score: 10812

Sequence: 1 MSGGAAGKSGSTPESLPLSP.....NISNLQKSGISNPPGSLNLRIT 2136

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Deiop 6.0 , Deilext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2/USPTO.spool/US10010720/runat.23092004.154319.25242/app.query.fasta.1.2311
-DB=GenEmbl -OEMT=fastap -SUFFIX=rge -MINMATCH=0 -LOOFCU=0 -LOOFEVT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10010720 @CGM 1.1 8347 @runat.23092004.154319.25242 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEIOBRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_htg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*
15: em_ba:.*
16: em_fun:.*
17: em_hum:.*
18: em_in:.*
19: em_mu:.*
20: em_om:.*
21: em_or:.*
22: em_ov:.*
23: em_pat:.*
24: em_ph:.*
25: em_pl:.*
26: em_ro:.*
27: em_sts:.*
28: em_un:.*

29: em_vi:.*
30: em_htg_hum:.*
31: em_htg_iny:.*
32: em_htg_other:.*
33: em_htg_mus:.*
34: em_htg_pln:.*
35: em_htg_rdi:.*
36: em_htg_mam:.*
37: em_htg_vtc:.*
38: em_sy:.*
39: em_htgo_hum:.*
40: em_htgo_mus:.*
41: em_htgo_other:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10810	100.0	6411	6 AR302594	AR302594 Sequence
2	10794.5	99.8	7280	6 AX766337	AX766337 Sequence
3	10753.5	99.5	6690	6 AR302591	AR302591 Sequence
4	10677	98.8	7149	6 AR302588	AR302588 Sequence
5	10677	98.8	7149	6 HSA296280	AJ296280 Homo sapi
6	10646	98.5	6327	6 AR302602	AR302602 Sequence
7	10589.5	97.9	6606	6 AR302599	AR302599 Sequence
8	10513	97.2	7065	6 AR302596	AR302596 Sequence
9	10508	97.2	6231	6 AR302610	AR302610 Sequence
10	10451.5	96.7	6510	6 AR302607	AR302607 Sequence
11	10435	96.5	6812	6 AB002342	AB002342 Human mRN
12	10375	96.0	6969	6 AR302604	AR302604 Sequence
13	10344	95.7	6147	6 AR302618	AR302618 Sequence
14	10287.5	95.1	6426	6 AR302615	AR302615 Sequence
15	10211	94.4	6885	6 AR302612	AR302612 Sequence
16	9929	91.8	6000	6 AR302595	AR302595 Sequence
17	9872.5	91.3	6279	6 AR302592	AR302592 Sequence
18	9796	90.3	6918	6 AR302589	AR302589 Sequence
19	9765	90.6	5736	6 AR302603	AR302603 Sequence
20	9729.5	90.0	7328	6 AX056418	AX056418 Sequence
21	9708.5	89.8	6195	6 AR302600	AR302600 Sequence
22	9632	89.1	6654	6 AR302597	AR302597 Sequence
23	9627	89.0	5820	6 AR302611	AR302611 Sequence
24	9570.5	88.5	6099	6 AR302608	AR302608 Sequence
25	9494	87.8	6558	6 AR302605	AR302605 Sequence
26	9463	87.5	5736	6 AR302619	AR302619 Sequence
27	9406.5	87.0	6015	6 AR302616	AR302616 Sequence
28	9330	86.3	6474	6 AR302613	AR302613 Sequence
29	9204	85.1	7342	10 AF227741	AF227741 Rattus no
30	9041.5	83.6	9313	10 AY309076	AY309076 Mus muscu
31	4053	37.5	2949	6 AR302590	AR302590 Sequence
32	4043	37.4	2490	6 AR302593	AR302593 Sequence
33	3889	36.0	2865	6 AR302598	AR302598 Sequence
34	3879	35.9	2406	6 AR302601	AR302601 Sequence
35	3783.5	35.0	5061	10 BC056761	BC056761 Mus muscu
36	3751	34.7	2769	6 AR302606	AR302606 Sequence
37	3741	34.6	2310	6 AR302609	AR302609 Sequence
38	3587	33.2	2685	6 AR302614	AR302614 Sequence
39	3577	33.1	2226	6 AR302617	AR302617 Sequence
40	3408	31.5	2653	9 AC004803	AC004803 Homo sapi
41	3107	28.7	154369	9 AX224728	AX224728 Homo sapi
42	2953	27.3	2310	6 BC055788	BC055788 Mus muscu
43	2844	26.3	2640	10 HSA242724	AJ242724 Homo sapi
44	2638.5	24.4	6891	9 AB051547	AB051547 Homo sapi
45	2621	24.2	7792	9	

RESULT 1

ALIGNMENTS

AR302594 AR302594 6411 bp DNA linear PAT 12-JUN-2003
 LOCUS AR302594
 DEFINITION Sequence 13 from patent US 6541252.
 ACCESSION AR302594
 VERSION AR302594.1 GI:31690892
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 6411)
 Walke, D.W., Hilbun, E., Donoho, G. and Turner, C.A. Jr.
 TITLE Human kinases and polynucleotides encoding the same
 JOURNAL Patent: US 6541252-A 13 01-APR-2003;
 FEATURES
 Location/Qualifiers
 1..6411
 /organism="unknown"
 /mol_type="genomic DNA"

Alignment Scores:
 Pred. No.: 5.52e-258 Length: 6411
 Score: 10810.00 Matches: 2136
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.98% Indels: 0
 DB: Gaps: 0

US-10-010-720-14 (1-2136) x AR302594 (1-6411)

QY 1 MetSerGIyGIyAlaIaIaGIyGSIySerSerTherProGIySerLeuPheLeuSe-Pro 20
 Db 1 ATGTTCTGCGCGCGCGGAGAGAGACAGACAGACTCCCGTTCTCTCTCTCTCGCG 60
 QY 21 ProIaIaProIaIaProIaIaSerSerSerSerSerSerSerValGIyGIyLLeuGIy 40
 Db 61 CGCGCTCTCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 41 AlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 60
 Db 121 GCCCGCGCGCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 QY 61 MetAspLysAspSerSerGIyAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 80
 Db 181 ATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 QY 81 ArgThrSerValIleCyAspSerSerSerSerSerSerSerValIaIaIaIaIaIaIaIa 100
 Db 241 CGCGGAGCGGATCTGGAATCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 300
 QY 101 SerLeuProGIyProSerSerIleProIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 120
 Db 301 TCCCTGCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 QY 121 ArgGIyGIyThrValIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 140
 Db 361 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 141 AlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 160
 Db 421 GCCCGCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 161 SerLysAspAArgProValSerGIyProSerSerLeuValGIySerLysGIyGIyGIy 180
 Db 481 AGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 QY 181 AlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 200
 Db 541 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 QY 201 GlnAspAspIleGIyGIyLLeuGIyThrLysAlaValGIyMetSerSerAspGIyAArgPhe 220
 Db 601 CAGATGATATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

QY 221 LeuLysPheAspIleGIyGIyLLeuGIyAArgGIySerPheLysTherValIyLysGIyLeuAsp 240
 Db 661 CTCAGATTGACATCGAAGATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 QY 241 ThrGIyThrValGIyAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 260
 Db 721 ACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 QY 261 GluArgGlnAArgPheLysGIyGIyAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 280
 Db 781 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 QY 281 ValAArgPheTyraSerSerTrpGIySerTherValIyGIyLysLysCysIleValLeuVal 300
 Db 841 GTTAGATTATATATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 QY 301 ThrGluLeuMetTherSerGIyThrLeuLysThrTyLeuLysAArgPheLysValMetLys 320
 Db 901 ACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 QY 321 IleLysValLeuAArgSerTrpCysAArgGlnIleLeuLysGIyLeuGlnPheLeuHisThr 340
 Db 961 ATCAAGATTCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 QY 341 ArgThrProProIleIleHisAArgPheLysCysAspAsnIlePheIleThrGIyPro 360
 Db 1021 CGAATCTCACTTATCATTCACCGAGATCTTAATGTACAAACATCTTATCACCGGCGCT 1080
 QY 361 ThrGlySerValIleGIyLysPheLysGIyLeuAlaIaIaIaIaIaIaIaIaIaIaIaIa 380
 Db 1081 ACTGAGCTCACTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 QY 381 LysSerValIleGIyThrProGIyPheMetAlaProGIyMetTrpGIyGIyLysTyraSp 400
 Db 1141 AAGAGTGTATAGATCCCAAGAGTGTATGCGCTGAGATGTATGAGAGAGAGAGAGAT 1200
 QY 401 GluSerValAspValTyraIaPheGIyMetCysMetLeuGluMetAlaIaIaIaIaIaIa 420
 Db 1201 GAATCCCTTACGTTATATGCTTTGGAGATGATGATGATGATGATGATGATGATGATGAT 1260
 QY 421 ProTySerGIyCysGlnAsnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 440
 Db 1261 CTTACTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 QY 441 ProIaIaSerPheAspLysValAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 460
 Db 1321 CCAGCCAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 QY 461 ArgGlnAsnLysAspGIyAArgTyraSerIleLysAspLeuLeuAsnIaIaIaIaIaIa 480
 Db 1381 CGACAAACAAAGATGAAGATATTCATCAAGACCTTTGAACCATGCTTCTTCCAA 1440
 QY 481 GluGIyThrGIyAlaArgValGluLeuAlaGIyLysAspGIyGIyLysIleAlaIle 500
 Db 1441 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 QY 501 LysLeuThrLeuAArgIleGIyAspIleLysLysLeuLysGIyLysTyraLysAspAsnIu 520
 Db 1501 AAATTATGCTACATATTAAGATATTAAGAAATTAAGAAATTAAGAAATTAAGATATTA 1560
 QY 521 AlaIleGIyPheSerPheAspLeuGIyAArgAspValProGIyAspValAlaIleGluMet 540
 Db 1561 GCTATTGAGTTTCTTTTATTTAGAGAGAGATGCCAGAGATGTTCACAGAAATG 1620
 QY 541 ValGIySerGIyTyraValCysGIyLysAspHisLysThrMetAlaIaIaIaIaIaIa 560
 Db 1621 GTAGAGTCTGGATGTCTGTGAAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
 QY 561 ArgValSerLeuIleLysAArgLysAArgGIyLysArgGlnMetValAArgGIyGIyGIy 580
 Db 1681 AGAGATCATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
 QY 581 LysLysGIyGIyGIyLysSerSerLeuLysGIyGlnValGIyGlnSerSerAlaSerGln 600

Db 6121 GAACCTGGGCAACTGTCATCTCCAGACTCCGAACTCGGTGGCTCTGCCCCCATCTCT 6180
 QY 2061 AAlaAsSerAlaThSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnTyr 2080
 Db 6181 GAGAGATGAGTACCTCTCTAGCTCATTACCAACTATATGTCGCCCCACAGCAAGTAT 6240
 QY 2081 GlyPheProAlaThrProPheGlyAlaGlnTyrPheSerGlyThrGlyGlyProAlaProGln 2100
 Db 6241 GGCTTCCAGCTACCCCATTTGGCGCTCATATGAGAGGAGCGGTGGCCACAGCACAG 6300
 QY 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheGlnLysSer 2120
 Db 6301 CCACTGGCCAGTTCACACTCTGGGAATGCTCTCTTGCAAAATTTCAACATCAGCAAT 6360
 QY 2121 LeuGlnLysSerLysSerAsnProProGlySerAsnLeuArgThrThr 2136
 Db 6361 TTGCAGAAATCATCAGCAACCCCGAGCTCCAACTCGGAGCACT 6408

RESULT 2

AX766337

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 Gurturjan, R., Hafalia, A.J., Patterson, C., Tribouley, C.M.,
 Azimzai, Y., Ding, L., Yao, M.G., Yue, H., Au-Yang, J., Bandman, O.,
 Baughn, M.R., Butford, N., Burrill, J.D., Elliott, V.S., Gandhi, A.R.,
 Ison, C.H., Kearney, L., Lal, P.G., Lu, D.A., Lu, Y., Tang, Y.T.,
 Zhangler, K.A., Marcus, G.A., Nguyen, D.B., Policky, J.L., Ramkumar, J.,
 Thangavelu, K., Thornton, M., Wallis, N.K. and Warren, B.A.
 Human kinases
 Patent: WO 0218557-A 33 07-MAR-2002;
 Incyte Genomics Inc. (US)
 Location/Qualifiers
 source 1..7280
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="ID No: 5502218CB1"

ORIGIN

Alignment Scores:

Pred. No.: 1 53e-257 Length: 7280
 Score: 10794.50 Matches: 2134
 Percent Similarity: 99.91% Conservative: 0
 Best Local Similarity: 99.91% Mismatches: 1
 Query Match: 99.84% Indels: 1
 Gaps: 1

US-10-010-720-14 (1-2136) x AX766337 (1-7280)

QY 1 MetSerGlyGlyAlaAlaGlnLysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 419 ATGTCTGGCGGCGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 478
 QY 21 ProAlaProAlaProLysAsnGlySerSerSerSerSerValGlyGlnLysLeuGly 40
 Db 479 CCGGCTCCGCGCCCGCAAGATGCTCCAGCTCCGATTCCTCGGTGGGAGAACTGGGA 538
 QY 41 AAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 60
 Db 539 GCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 598
 QY 61 MetAspLysAspSerAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 80
 Db 599 ATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658

QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100
 Db 659 CCGCGGAGAGGATATCTGAGCTTCAATGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 718
 QY 101 SerLeuProGlnProSerLysProAlaAlaValAlaProGlnSerAlaProProGlyProHis 120
 Db 719 TCCCTGCCCGAGCCAGCATCCCGCGGCTGTCGCCAGAGTCTCCACCGAGCGCCAC 778
 QY 121 ArgGlnGlnThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 779 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 838
 QY 141 AAlaAlaProGlyGlnGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 160
 Db 839 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 898
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGlyProProPro 180
 Db 899 AGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 958
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGlnProGlnGlnGlnAlaArgSerGln 200
 Db 959 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1018
 QY 201 GlnAspAspIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
 Db 1019 CAGGATATATCGAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1078
 QY 221 LeuLysPheAspIleGlnIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
 Db 1079 CTCAAGTTTGAATCGAATCGAATCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1138
 QY 241 ThrGlnThrThrValGlnValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 260
 Db 1139 ACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1198
 QY 261 GlnArgGlnArgPheLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 Db 1199 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1258
 QY 281 ValArgPheLysAspSerTrpGlnSerThrValLysGlyLysLysCysIleValLeuVal 300
 Db 1259 GTTAGATTTTATGATCTCTGGAATCCAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1318
 QY 301 ThrGlnLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
 Db 1319 ACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1378
 QY 321 IleLysValLeuArgSerTrpCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
 Db 1379 ATCAAGGTTCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1438
 QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 Db 1439 CGAAGCTCCAGCTATCTTCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1498
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
 Db 1499 ACTGCTCAGTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1558
 QY 381 LysSerValIleGlyThrProGlnPheMetAlaProGlnMetTyrGlnLysTyrAsp 400
 Db 1559 AAGAGTGTATAGTACCCCAAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1618
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlyTyr 420
 Db 1619 GAATCGTTGAAGTTATGCTTTTGGATGTGATGCTTGAATGCTTGAATGCTTGAATGCTT 1678
 QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
 Db 1679 CCTTACTCGAGATGCGCAAAATGCTGCGAGATCTACGTCGCTGCGTGCAGTGGGTGAAG 1738

QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluLysCysIle 460
DB 1739 CCAGCCAGATTTTGCACAAAGTAGCAATTCCTGAGTAGAAGAAATTTATGAAAGATGCATA 1798
QY 461 ArgGlnAsnLysAspGluArgLysSerIleLysAspLeuAsnHisAlaPhePheGln 480
DB 1799 CGACAAAACAAAGATGAAAGATATTCATCCAAAGACCTTTTGAACCATGCTTCCCA 1858
QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
DB 1859 GAGGAAACAGAGATACGGGTAGAAATTTAGCAAGAAAGATGATGAGAAAAAATAGCCATA 1918
QY 501 LysLeuThrLeuArgIleGluAspLleLysLysLeuLysGlyLysThrLysAspAsnGlu 520
DB 1919 AAATTATAGCTACGATTTGAGATTTAAAGAAATTTAAAGGAAATTCAAAGATATGAA 1978
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGluLysMet 540
DB 1979 GCTATTAGATTTTCTTTTGAATTTAGAGAGATGCTCCCAAGAGATGTTGCACAAAGAAATG 2038
QY 541 ValGluSerGlyThrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
DB 2039 GTAGAGTCTGGGTATGCTCTGAGAGTGATCACAAGACCATGGCTAAAGCTATCAAGAAC 2098
QY 561 ArgValSerLeuLleLysArgLysArgGluGluArgGluLeuValArgGluGluGlu 580
DB 2099 AGAGTATCATTAATTAAGAGAAACGAGAGCGGAGTTGGTAGGAGAGACAAAGAA 2158
QY 581 LysLysLysGluGluGluSerSerLeuLysGluGluValGluGluSerSerAlaSerGln 600
DB 2159 AAAAAAAGCAGAGAGAGAGAGAGCTCTCAACAGCAGGTAGAAACAATCCAGTCTCCAG 2218
QY 601 ThrGlyLleLysGluLeuProSerAlaSerThrGlyLleProThrAlaSerThrThrSer 620
DB 2219 ACAGGAATTCAGAGAGCTCTCTGCTAGCACCAGCATACCTACTGCTTCTACCACTTCA 2278
QY 621 AlaserValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGluGluLeu 640
DB 2279 GCTTCAGATTTTACACAAAGTAGAACCTTAAGAACTTGAGGACATCAACATCAACACTA 2338
QY 641 GlnThrGluGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
DB 2339 CAGTACCAAGCAACCAAGTATATCTGTATCTGATGGAGCGTTGACAGTGTCAAGGAA 2398
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerThrGlySerGln 680
DB 2399 TCCTGTGCTTCCACAGAACTCTGAGTAGAGCCCAACAGACAGTTCATATGGTTCCTCAA 2458
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGluHisIleProSerThrValGlnAla 700
DB 2459 CATGAAACAGGCACTTTCACAGGCAAGCCCAAGGCAATATCCTTCTACTGCTCCAAAGCA 2518
QY 701 GlnSerGlnProHisGlyValThrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
DB 2519 CAGTCTCAGCCCAAGGGGTATATCCACCTCAAGTGTGGCAAGGGGAGAGCCAGGCT 2578
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
DB 2579 CAGCCATCTCTCAAGTAGCTTAACAGGGGTTTCATCTTCCCAACCAATACATACCTCT 2635
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnThrSerLeuSer 760
DB 2636 CAGCAGCAGGAAATACAGAGACAGCCCTCTCCCAACAGACAGTGCAGATTCACCTTCA 2695
QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
DB 2696 CAGACATCAACCTCCAGTAGGCCACTACTGCACACCAAGTAGTCAACCTCAAGCTCCA 2755
QY 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
DB 2756 CAAGTCTTGCTCAAGTATCATGCTGAGAAAAACAGAGTACTCAGGAGATCTCTCAGGTGCT 2815
QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820

DB 2816 CCTGCAGACCGCAGTTTGCAGTACACAGCCCAAGCTAACCCAGCCGACCACTTGGCTCTCC 2875
QY 821 SerValAspSerAlaHisSerSerPylAlaSerGlyMetSerAspLysGlnGluAlaVal 840
DB 2876 TCTGTAGACAGTGCACATTCAGATGTTGCTTCAAGATATGATGATGATGATGATGATGAT 2935
QY 841 ProSerSerSerGlyArgHisGluGlyArgThrThrLysArgHisThrArgLysSerVal 860
DB 2936 CCATCTTCCAGTGAAGGATGAGAGAAAGAACTACAAAGCGCATTAACCAAAATCTGTA 2995
QY 861 ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArgLleLeuAsnValSer 880
DB 2996 AGAGTCTGCTCGACATGAAAAAACTTCAAGCCCAAAATTAAGAAATTTGAATGTGTCA 3055
QY 881 AsnLysGlyAspArgValValGluCysGlnLeuGluIleThrHisAsnArgLysMetValThr 900
DB 3056 AATTAAGAGACCGAGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3115
QY 901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn 920
DB 3116 TTCAAAATTTGACCTAGATGTTGACCAACCCAGAGATAGCAACAAATTAATGATGATCA 3175
QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
DB 3176 GACTTATCTTACGAAATAGAGAGAGAGAGTCTTGTGATCAAGTGCAGAAATTTATTGAA 3235
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGlyAspGlnGlyLeu 960
DB 3236 AAGCTATGATAAATGCTCAAGTAGAGATGATGATGATGATGATGATGATGATGATGATG 3295
QY 961 GluSerLeuGlnGlyLysAspAspThrGlyPheSerGlySerGlnLysLeuGluGlu 980
DB 3296 GAGAGTCTACAGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3355
QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyLleProThrSerSer 1000
DB 3356 TTCAAACAAACCAATTCCTGCTCTTCCATGACAGCAAGCAAAATGACATTCCTACCAAGTCT 3415
QY 1001 LeuThrGlnValAlaHisSerAlaGlyArgArgPheIleValSerProValProGluSer 1020
DB 3416 TTAACCTCAGTGTCTTCACTTCTGCGGAAAGCGGTTTATAGAGTCCGTGCCAGAAAGC 3475
QY 1021 ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
DB 3476 CGATTACAGAGATCAAAAGTTTCCCGCAGTGAATTAACAGATACAGTGTGCCCTCTCA 3535
QY 1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAla 1060
DB 3536 GCTCAGAGCCCTGGAAATGAACTTGTCTCACTGATCATCTCTTAAGTACACAGAGCC 3595
QY 1061 PheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe 1080
DB 3596 TTTTCTGAATTCAGCGTGCCTCAATATGACAGAAAGCAACATACAGACCTCAAACTTT 3655
QY 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal 1100
DB 3656 AGTATACAGAGCAAACTTTCAGTAGTACTCTCTTCTTAAGTAGCATTCCTGAGATC 3715
QY 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
DB 3716 CCACACACAGCAGCAGCAGCAGACACACACCTCTGCAACAGAGCCCTCTTAATGACAT 3775
QY 1121 SerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGlyVal 1140
DB 3776 TCCACATCAAGTATTCAGTCTAGGTACAGTCCACAGAGAGAGGATTCCTGAGATT 3835
QY 1141 AlaThrSerThrGlyValValThrSerGlyLysLeuProIleProProValSerGluSer 1160
DB 3836 GCCACACAGCAGAGGTGTGTAATTCAGAGTGTCTCCCATACCAACCTGTGTCTGATATA 3895
QY 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180

Db	3896	CCAACTACTTCCAGCGGATGTTTCAAGTATCAACAATACCTGCAGATTGCTCAATATCTACT	3955
QY	1181	ThiSerProSerLeuGlnValProThrSerThiSerGluIleValValSerSerThrAla	1200
Db	3956	ACATCCCGGTCATTCAAAGTCCCAATCCCAATCCACTGCAATGAGTGGTGTTCATGACACA	4015
QY	1201	LeuTyProSerValThrValSerAlaThrSerAlaSerAlaGlyGlySerThrAlaThr	1220
Db	4016	CTGATTCCTTCAGTAAACAGTTTCAGCACTTCAGCTCGCGGGGGGAGTACTGCTACC	4075
QY	1221	ProGlyProIysProhAlaValAlaSerGlnGlnAlaGlySerThrValGly	1240
Db	4076	CCAGGTCCTTAGGCTCCAGCTGTAAGTATCCAGCGGCGAGCGAGCACTACTGTGGGA	4135
QY	1241	AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer	1260
Db	4136	GGCAACATTAACATCAATTTCTACACCACTTATTCGCCAGAGCAAGCTTCACACTGACC	4195
QY	1261	IleGlnLeuSerSerSerThiSerThrProThrLeuAlaGluThrValValSerAla	1280
Db	4196	ATTCACTTAGAGCGAGCGACTTCTACTCTCTTAATTAGCTGAACCGGTGTAGTAGGCA	4255
QY	1281	HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla	1300
Db	4256	CATCTACTGATTAAGACATCTCATAGCAATAACATGAAATGGCTTCTCCCTCTGCA	4315
QY	1301	ProSerSerSerSerSerProGlyAlaGlyValSerSerTyIleSerGlnProGlyGly	1320
Db	4316	CCATCTTCTCTTCCCTCTCTGGACAGAGAGTCTAATTATTTCTCAAGCTGGGGG	4375
QY	1321	LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla	1340
Db	4376	CTGCATCCTTTGGTCAATTCACATGATGATGCTTACTCTTAATTTCCCAAGCAGCA	4435
QY	1341	GlyProThrSerThrProLeuLeuProGlnAlaProSerIleProProLeuValGlnPro	1360
Db	4436	GGACTACTTCTTACACCTTATTAATCCCAATACCTTGATATCCCACTTGATACAGCTTGG	4495
QY	1361	ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu	1380
Db	4496	GTTGCCAATGTCCTGCTGTACAGAGACACTAATTATATGTCAGCTTCAACAGAGCTTGG	4555
QY	1381	LeuProAsnGlnProHisThrHisCysProGlyValAspSerAspThrGlnProIysAla	1400
Db	4556	CTTCCCAACAGCCCCCACTACACTTGCTGCGAAGTATCTGTATACAAACCAACAGCT	4615
QY	1401	ProGlyIleAspAspIleIysThrLeuGlnGlnIleIysLeuArgSerLeuPheSerGluHis	1420
Db	4616	CCTGGAAATTGATGACATTAAGACCTTGAAGAAAAAGCTGGGGTCTGTTCACATGAAACAC	4675
QY	1421	SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer	1440
Db	4676	AGCTCATCTGGAGCTCAGCAATGCTCTGTCTCACTGAGAGCTCACTGATATAGAGACC	4735
QY	1441	ThrValThrProGlyIleProThrThrAlaValAlaProSerIysLeuLeuThrSerThr	1460
Db	4736	ACTGTCACACAGGACATCCAACTACAGCTGGTGGACCAAGCAAACTCTGACTCTACC	4795
QY	1461	ThiSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr	1480
Db	4796	ACAAATACCTGCTTACACCAACCAATTTTACACTGAGAAACAATTCCTTGGCAGTTTACA	4855
QY	1481	ProValValThrProGlyGlnValSerThrProValSerThrThrThiSerGlyValIys	1500
Db	4856	CCAATGTCTCACTCTGGGCAAGTTTCTACACCCAGTACAGACATTCATCAGAGAGTAAA	4915
QY	1501	ProGlyThrAlaProSerIysProProLeuThrIleAlaProValLeuProValGlyThr	1520
Db	4916	CTTGGAATCTCTCCCTTCCAAAGCACTCTAACTTAAGAGCTCCGAGTCTGCACTGGTACT	4975
QY	1521	GluLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu	1540
Db	4976	GAACTTCACAGACTCTACCCAGAGAGACAGCTGCAACTTTTCCAGGAACTTCTCTCA	5035

QY	1541	ThrnInserGlnInProIeugInuAspLeuAspAlaGlnLeuArgAyrThrIleuSerPro	1560
Db	5036	ACCCAGTCCCAACCACTCTTAAGAGATCTTGAAGCTCAATTGAGAAACAACCTAGTCCA	5095
QY	1561	Glu***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle	1580
Db	5096	GAGATGATCAAGTACCTTCTCGGTGGTGGTCCGTCCATGGCGGCTCCAAACAGCATC	5155
QY	1581	ThrnGluAlaGlyThrGlnProGlnuInuSglValSerGlnValIysGluGlyProValLeu	1600
Db	5156	ACAGAAGCAGAAACACAGCTTCAGAAAGGTGTTTCTCAAGTCAAAAGAGCCCTGTCTCA	5215
QY	1601	AlaThrSerSerGlyAlaGlyValPheIysMetGlyArgPheGlnValSerValAlaAla	1620
Db	5216	GCACCTACTTCAGAGCTGGTGGTCTTTAAAGATGGAGACGATTCAGGTTCTGTTCAGCA	5275
QY	1621	AspGlyAlaGlnIlySgluGlyIlyAsnIlySerGluuAspAlaIlySerValHisPheGlu	1640
Db	5276	GACGGTGGCCAGAAAGAGGTAAATAATGATCAGAAAGTCAAGCTGTTCAATTTGAA	5335
QY	1641	SerSerThrSerGluSerSerValIleuSerSerSerSerProGluSerThrIleuValIlys	1660
Db	5336	TCCAGCACTCAGAGTCTCTCAGTGTATCAAGTACTAGTCCAGAGAGTACCTTGTGAAA	5395
QY	1661	ProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGluSerAlaHis	1680
Db	5396	CCAAAGCCGAATGGCATTAACATCCCTGGTATCTCTTCAGATGTCCAGAGAGTGGCCAC	5455
QY	1681	LysThrThrAlaSerGluAlaIlySerAspThrGlyGlnProThrIlyValGlyArgPhe	1700
Db	5456	AAAACCTACTGCTCAGAGGCAAGTCAAGACACTGGGAGCCTTACCAAGTTGGAGCTTTT	5515
QY	1701	GlnValThrThrThrAlaAsnIlyValGlyArgPheSerValSerIlyThrGluuAspIlys	1720
Db	5516	CAGGTGCACTACAGCAACAAGTGGTGCTTCTCTGATCAAAAATCGAGGACAG	5575
QY	1721	IleThrAspThrIlySlySgluGlyProValAlaSerProPheMetAspLeuGlnIn	1740
Db	5576	ATCACTGACCAAAAGAAAGAGCCAGTGGCATCTCCTCTTATGATTTGGAACTA	5635
QY	1741	AlaValIleuProAlaValIleProIlySlyGlnIlyProGluIleuSerGlnProSerHis	1760
Db	5636	GCCTGTTCTTCTGCTGTGATACCAAAAGAAAGAGGCTGAACGTCAAGCCTTCACAT	5695
QY	1761	LeuAsnGlyProSerSerAspProGluAlaAlaPheIleuSerArgAspValAspAspGly	1780
Db	5696	CTAAATGGGCGGCTTCTGACCCGGAGCCGCTTTTAAAGTAGGAGTGTGGATGATGT	5755
QY	1781	SerGlySerProHisSerProHisGlnIleuSerSerIlySerIleuProSerGlnAsnLeu	1800
Db	5756	TCCGGTAGTCCACACTGCGCCCATCACTGAGCTGATCAAAAGAGCCTTCTTAGCGAATCTA	5815
QY	1801	SerGlnSerIleuSerAsnSerPheAsnSerSerTyrMetSerSerAspAsnGluSerAsp	1820
Db	5816	AGTCNAAGCTTAGTAATTCATTAATTAATCTCTTCACTAGTGAGCGCAATGAGTCAAT	5875
QY	1821	IleGluAspGluAspLeuIlySleuGlnLeuArgArgLeuArgAspIlySHisIleuIlySglu	1840
Db	5876	ATCGAAGATCAAGACTTAAGATTAGAGCTGCCAGCACTAGCAATAAACTTCCAAAGAG	5935
QY	1841	IleGlnAspLeuGlnSerArgIlySHisIleuIleGluSerLeuTyrThrIlySleuGly	1860
Db	5936	ATTCAAGACTTCAGAGTGGCGCAGAAAGCATGAATTAATCTTGTGATACCAACTGGGC	5995
QY	1861	IlyValProProAlaValIleIleProProAlaAlaProIleuSerGlyArgArgArg	1880
Db	5996	AAGGTGCCCCCTGCTGTATTAATTCCTCCAGCTGCTCCCTTTCAGAGGAAACACACGA	6055
QY	1881	ProThrIlySerIlySglIlySerIlySerSerArgSerSerIleuGlyAsnIlySerPro	1900
Db	6056	CCACTAAAGCAAGAGGACGAACATCTAGTCAGACAGTTCCTTGGGGATATAAAGCCCC	6115


```

QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 1021 CGAAGCTCCAGCTTATCATTCACCGCAATCTTAAATGAGACACATCTTATATCCCGGCGCT 1080
QY 361 ThrGlySerValIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
Db 1081 ACTGGCTCAGTCAGATGGATTGGAGACCTCGCTCGGCAACCTTAAACGGGCTCTCTTGGCC 1140
QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGlyLysTyrAsp 400
Db 1141 AAGAGTGTGATGAGTACCCAGAGTTCATGCGCCCTCGAGATGTATGAGGAGAAATATGAT 1200
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGlyTyr 420
Db 1201 GAATCCGTTGACGTTTATGCTTTTGGATGTGATGATGCTTGAAGTGGTAACTTCAATAT 1260
QY 421 ProTyrSerGluCysGlnAsnAlaIleGlnIleTyrArgValThrSerGlyValLys 440
Db 1261 CCTTACTCGAGTGGCAAAATGCTGCAAGATCTTACCTCGCTGACCAAGTGGGTTGAG 1320
QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
Db 1321 CCAGCAGCTTTTGACAAAGTACATCTCGAAGTGAAGAAATATATGAGAGATGACATA 1380
QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnIleAlaPhePheGln 480
Db 1381 CGAAGAAACAAAGATGAAAGATATTCATCAAGACCTTTTGAACCAATGCTTTTCCAA 1440
QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluLysAspGlyGlyLysIleAlaIle 500
Db 1441 GAGGAAACAGAGATGACGGTATGATTTAGCAGAGAAAGATGATGAGAAAGAAATATGCA 1500
QY 501 LysLeuThrLeuAlaGlyIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
Db 1501 AAATATAGGCTACCTATTTGAAGATATTAAGAAATTAAGGGAATAATCAAGATATAGAA 1560
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 1561 GCTATTGAGTTTCTTTTGAATTAGAGAGAGATCTCCAGAAAGTGTGGACAAAGAAATG 1620
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 1621 GTAGAGTCTGGATATGCTGTGAAGGTGATCAAGACCAATGGCTTAAAGCTTATCAAGAC 1680
QY 561 ArgValSerLeuIleLysArgLysArgGluArgGlnArgGlnLeuValArgGluGluGln 580
Db 1681 AGAGTATCATTAATTAAGAGAAACGAGACGAGGAGATTGTGAACGGAGGAGCAAGAA 1740
QY 581 LysLysLysGlnGluGluSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln 600
Db 1741 AAAAAAAGCAGAGAGAGACAGCTCTCAACACAGAGTAAACAAATCCAGTCTCCAG 1800
QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1801 ACAGGATCAAGCAGCTCCCTTCTGCTAGCACCGGCACTTACTGCTTCTACCACTTCA 1860
QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db 1861 GCTTCAGTTTCAACAGATGAACCTGAAGAACTGAGGACAGATCAACATAACAACTA 1920
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1921 CAGTACCAAGAACCCAGTATATCTGTTATCTGATGGAGACGGTTTACAGTGTCTCAGGGA 1980
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1981 TCCTCTGCTTCAACAGATCTGAGTGAACAGCAGCAACACAGATTTCATATGTTCCAA 2040
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2041 CATGAACAGCAGCATTTCTACAGGCAACAGCCAGGCAATATACCTTCTACTGCCAAGCA 2100
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720

```

```

Db 2101 CAGTTCAGCCCAATGGGATATATCCACCTTCAATGTGGACAGAGGAGCAGCCAGGCT 2160
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2161 CAGCCATCTCAAGATGCTTAAACAGGGGTTTCATCTTCCCAACCATTAACATCTCAG 2220
QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2221 CAGCAGCAGGGAATACAGACAGACAGCCCTCTCAACAGACAGTGCAGATTCACCTTCA 2280
QY 761 GlnThrSerThrSerSerSerGluAlaThrThrAlaGlnProValSerGlnProGluAlaPro 780
Db 2281 CAGCATCAACCTCCAGTGAAGGCCATCTACTCAACGCCAGTGAAGTCAAGCTTCAAGCTCA 2340
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 2341 CAAGTCTTGCTCAAGTATCAGTGAAGAAACAGGGGCTTCCCACTCGACCTGCACACAG 2400
QY 791 ----- 791
Db 2401 TACCCAGAGATTAATAATTTGCTCCTCTTCCAAAGTGGCTTCTGTTGATCCATCT 2460
QY 791 ----- 791
Db 2461 ACAAGTCTTAMCCCTCCCATGCGCAGAGATAGTACTGCTACCTGGGACTTTCACCA 2520
QY 791 ----- 791
Db 2521 GTGGTGAGCTTATGTGAATCAATCTTTAGTCTTATGGGTGTGTAGAGAGACAG 2580
QY 791 ----- 791
Db 2581 GTTCAAGTGTCCCAAGCAGAGGAGTTAGCAACAGCCCACTACATCTCCAGCAA 2640
QY 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGluProValAlaVal 807
Db 2641 GCAGTTTGGAGATGACTCAGGAGTCTCTCAGGTGCTCTCTGACAGGCCAGTGTGACATA 2700
QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2701 GCACAGCCCAAGCTACCCAGCCGACCACTTGGCTTCTCTGTAGACATGACATTTCA 2760
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyArgHis 847
Db 2761 GATGTTGCTTCAAGTATGATGATGATGCAATGAAGAAAGTCCCATCTTCCAGTGAAGGAT 2820
QY 848 GluGlyArgThrThrLysArgHisTyrArgLysSerValArgSerArgSerArgHis 867
Db 2821 GAAGGAAGAACTACAAAGCGCATTTACGAAATCTGTAAAGAGTGGCTTCTGACATGAA 2880
QY 868 LysThrSerArgProLysLeuArgIleLeuAsnValSerAsnLysGlyAspArgValAla 887
Db 2881 AAAAATTCACGCCCAAAATTAAGAAATTTGAATGTTTCAAAATTAAGAGACCGAGTAGTA 2940
QY 888 GluCysGlnLeuGlnThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly 907
Db 2941 GAATGCTCAATTAGAGCTATATAGAAAAGTGTACATTAATTAATTAACCTTACATGATGT 3000
QY 908 AspAsnProGluGluIleAlaThrIleMetValAsnAspAspPheIleLeuAlaIleGlu 927
Db 3001 GACAACCCCGAGAGATGACCAACATTAATGAGAACAAAGATTTATTTCTAGCAATAGAG 3060
QY 928 ArgGluSerPheValAspGlnValArgGluIleIleGluLysAlaAspGluMetLeuSer 947
Db 3061 AAGAGTGTGTTGTGATCTCAAGTGCAGAAATTAATTAAGAAAGATGAATATGCTCACT 3120
QY 948 GluAspValSerValGluProGluLysAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
Db 3121 GAGGATGTCAAGTGTGAACCAAGGCTGATTCAGGGAATGGAGAGTCTCAAGAAAGAGAT 3180
QY 968 AspTyrGlyPheSerGlySerGlnLysLeuGluGluGluPheLysGlnProIleProAla 987

```


Db 3181 GACTATGCTTTTCAGGTTCTCAAAATTGAAAGGAGTTCAAACCACTTCCTGG 3240
QY 988 SerSerMetProGlnGlnIleGlyIleProThrSerSerLeuThrGlnValAlaHisSer 1007
Db 3241 TCTTCATGCGACAGCAAAATAGGCAATTCCTACAGGTTCTTTAACTCAAGTTGTTCACTTC 3300
QY 1008 AlaGlyArgArgPheIleValSerProValProGlnSerArgLeuArgGlnSerVal 1027
Db 3301 GCGGGAAGGGGCTTTTATGAGATGCTGCGCAAAAGCGAATACAGAAATCAAAAGTT 3360
QY 1028 PheProSerGlnIleThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
Db 3361 TTCCCCAGTAAATACAGATACAGTTGCTGCTCTACAGCTCAGAGCCCTGGAAAGAAC 3420
QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnAlaPheSerGlnLeuArgArgAla 1067
Db 3421 TTGTCTCAGCTGCATCATCCCTTAGCTCAACAGAGCCCTTTCTGAACTTAGACGTGCC 3480
QY 1068 GlnMetThrGlnGlyProAlaThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
Db 3481 CAAATGACAAAGAGACCAAAVACAGACCTCCAAACTTGTACATACAGAACCAACTTT 3540
QY 1088 ProValValProProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaAlaThr 1107
Db 3541 CCAAGTAGTACTCTCTTTCTTAAGTAGCATGTGAGAGTCCCAACCAAGCAGAGCCACA 3600
QY 1108 AlaProValProAlaThrSerSerProProAsnAspIleSerThrSerValIleGlnSer 1127
Db 3601 GCACAGATCCCTCGAACAGACAGCCCTCTATATGACATTCACATACATAGTAATCACTCT 3660
QY 1128 GlnValThrValProThrGlnGlnGlyIleAlaGlyValAlaThrSerThrGlyValVal 1147
Db 3661 GAGGTTCAGTGGCCACTGAAAGAGAGGATGCTGAGATTGCCACAGACAGAGGTGTGTA 3720
QY 1148 ThrSerGlyGlyLeuProIleProProValSerGlnSerProValLeuSerSerValVal 1167
Db 3721 ACTTCAGGTGCTCTCCCATACCACTGTGTCTGATACCAAGTACTTCCAGGTAGTT 3780
QY 1168 SerSerIleThrIleProAlaValValSerIleSerThrThrSerProSerLeuGlnVal 1187
Db 3781 TCAGATACCAAAATCTGAGTGTCTCAATATCTACTACATCCCTCCCTCACTCAACTC 3840
QY 1188 ProThrSerThrSerGlnIleValValSerSerThrAlaLeuIleProSerValThrVal 1207
Db 3841 CCCACATCCACATCTGAGATGTTGTTCTTAGTACAGACCTGTATCTCTCCAGTAAAGTT 3900
QY 1208 SerAlaThrSerAlaSerAlaGlyGlySerThrAlaThrProGlyProIleProProAla 1227
Db 3901 TCAGCACTTACAGCTCTGCAAGGGGAGTACTGCTACCCAGGTCCTAAGCTCCAGCT 3960
QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
Db 3961 GTAGATCTCAGCAGGAGCAGGAGCAGCACTACTGAGAGGCCAATTAACAATCAATTTCT 4020
QY 1248 ThrThrThrSerPhePheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThr 1267
Db 4021 ACCACCACTTCACTCCAGACACACACTTACAGCTGTTCATTCAGCTTACGACGAGTACT 4080
QY 1268 SerThrProThrLeuAlaGlnThrValValValSerAlaHisSerLeuAspIleThrSer 1287
Db 4081 TCTATCTCTACTTGTAGTGAACCGGTGATAGTACGACACTCACTAGATAGACATCT 4140
QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerSerPro 1307
Db 4141 CATACAGACAAACAGATGCTTCTCCCTCTCTGCAACATCTTCTCTCTCTCTCT 4200
QY 1308 GlyAlaGlyValSerSerTyrIleSerGlnProGlyGlyLeuHisProLeuValIlePro 1327
Db 4201 GAGAGGAGAGGTCTAGATTATTTCTACGCTGTGGGCTGCACTCTTGGTATTCACA 4260
QY 1328 SerValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
Db 4261 TCAGTGAATGCTTCTACTATTTCTTCCCAAGCAGAGACCTAATCTTACACTTTA 4320

QY 1348 LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal 1367
Db 4321 TTACCCCAAGTACTAGTATCCCACTTGTGACAGCTCTTGCCTGCTGTGA 4380
QY 1368 GlnGlnThrLeuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
Db 4381 CAGCAGACACTAATCATATGATGAGCTCTCAACAGAGCTTGTGCTTCCCAACAGGCCCACTACT 4440
QY 1388 HisCysProGlnValAspSerAspThrGlnProIleValAlaProGlyIleAspAspIleCys 1407
Db 4441 CATTGCTCGAAGTAAATTTCTGATACCAACCCAAAGCTCTGGAATTGATGACATTAAG 4500
QY 1408 ThrLeuGlnGlnLeuSerLeuArgSerLeuPheSerGlnHisSerSerSerGlyAlaGlnHis 1427
Db 4501 ACTCTGAGAGAAAGCTGCGGTCTCTGTTCAAGAACAGACGTCATCTGGAGCTCAGCAT 4560
QY 1428 AlaSerValSerLeuGlnThrSerLeuValIleGlnSerThrValThrProGlyIlePro 1447
Db 4561 GCTCTGTCTCAGTGAAGACCTCATATGATAGAGACCTGTCAACACAGGAGCATCCCA 4620
QY 1448 ThrThrAlaValAlaProSerIleLeuLeuThrSerThrThrCysLeuProPro 1467
Db 4621 ACTACTGCTGTGCACCAAGCAAACTCTGACTTCTACCAAGTACTTGTCTTACCAACA 4680
QY 1468 ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValValThrProGlyGln 1487
Db 4681 ACCAATTTACACATAGGAACAGTTGCTTTCAGATTACACAGAGGTGACACAGCTGGGCAA 4740
QY 1488 ValSerThrProValSerThrThrThrSerGlyValIleProGlyThrAlaProSerIle 1507
Db 4741 GTTTCATCCCAAGCAGACCTACTACATCAGAGGTGAACCTGGAATGCTCTCCCTCAGG 4800
QY 1508 ProProLeuThrIleValAlaProValAlaLeuProValGlyThrGlnLeuProAlaGlyThrLeu 1527
Db 4801 CCACCTTAATCTAAGGCTCCGAGTCTCCAGTGGGTACTGAATCTTCCAGCAGGTAATCTA 4860
QY 1528 ProSerGlnGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnGlnProLeu 1547
Db 4861 CCAGCAGAGAGAGTGCACCTTTCCAGGACCTTCTCAACCCAGTCCAGCAACTCTA 4920
QY 1548 GlnAspLeuAspAlaGlnLeuArgArgThrLeuSerProGlu***IleThrValThrSer 1567
Db 4921 GAGGATCTTGATGCTCAATTGAGAAACAATGATGTCAGAGATKATACAGTACTTCT 4980
QY 1568 AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGlnAlaGlyThrGlnPro 1587
Db 4981 GCGGTGTGCTGTGTCCATGCGGCTTCCAAACAGCAATCACAGAAAGCAAGCAAGCT 5040
QY 1588 GlnIleGlyValSerGlnValIleGlyProValLeuAlaThrSerSerGlyValAlaGly 1607
Db 5041 CAGAAAGGTGTTTCTCAAGTCAAAGAGGCGCTGCTCAGCAACTAGTTCAGAGACTGT 5100
QY 1608 ValPheIleMetGlyArgPheGlnValSerValAlaAlaAspGlyAlaGlyIleGly 1627
Db 5101 GTTTTAAAGATGGACGATTTTCAGGTTCTGTGAGAGACGGTGCCCAAGAAAGAGGT 5160
QY 1628 LysAsnLysSerGluAspAlaIleSerValHisPheGlnSerSerThrSerGlnSerSer 1647
Db 5161 AAAATTAAGTACAGAAAGATGCAAGTCTGTTCATTTTGAATCCAGCACTCAGAGTCTCA 5220
QY 1648 ValLeuSerSerSerSerProGlnSerThrLeuValIleProGlnProAsnGlyIleThr 1667
Db 5221 GTCATACAGATAGTACGACAGAGTCTTGTGAAACCAAGACCAAGCCAAATGGCAATACC 5280
QY 1668 IleProGlyIleSerSerAspValProGlnSerAlaHisIleThrThrAlaSerGlnAla 1687
Db 5281 ATCCCTGTATCTCTTCAGATGTGCCAAGATGCCCAAGAACTACTGCTTCCAGAGGCA 5340
QY 1688 LysSerAspThrGlnGlnProThrIleValGlyArgPheGlnValThrThrThrAlaAsn 1707
Db 5341 AAGTACAGACCTGGCAGCTTACCAAGTTGACGCTTTTCAAGTACAACTACAGAAAC 5400

QY	1708	LVSVALGIYARXPHSESRVALSERIYSTRHIGUASPLYSILETHASPTHILYSIGLU	1727
Db	5401	AAAGTGGGTCGGTTCTCTGTATCAAAACCTGAGGACCAAGATCACTGACAAAGAAAGAA	5460
QY	1728	GIYPRIVAIASERPROPBHEMETSAPLEUGIUGIHALVALLEPROLAVALL	1747
Db	5461	GGACCAAGTGCATCTTCCTCTTTNTGATTTGAAACCAAGCTTCTTCGTGTGTA	5520
QY	1748	PROLYALYSGLULYBPROGILUENBERGIUPROSETHI.SLEASNGIYPROSESERASP	1767
Db	5521	CCAAAGAAAGAGAGAGCCTTGAACTGTCAAGCCTTCAACATCTMAAGGGCCGCTTCTGAC	5580
QY	1768	PROGILUALALAPHELEUSERARGASPVALASPAAGIYSEGIYSEPROHISESERPRO	1787
Db	5581	CCGAGGGCCGCTTTTAAAGTAGGATGTGATGTATGTTCGGTAGTCCACACTCGCC	5640
QY	1788	HISGLINLEUSERSELYSERLEUPROBERGIANENLEUSERGINSELEUSERHANSER	1807
Db	5641	CATCAGCTGAGTCCAAAGAGCCTTCTTGACCAAGATCTTAAAGCTTAAAGCTTAAATCA	5700
QY	1808	PHASANSERSELYMETSESERASPAANGIUSERAPILEGILUASPGILUASPLEULYS	1827
Db	5701	TTTAACTCCTCTTACATAGTAGACGACATAGTACGATATCGAAGATGAAGACTTAAAG	5760
QY	1828	LEUGILULEAARGAGLEUAARGASPLYSISLEULYSIGUILEGIASPLEUGINSEARG	1847
Db	5761	TTAGAGCTGCGACGACTACGAGATMAACATCTCAAGAGATTCCAGACCTCAGAGCTGCG	5820
QY	1848	GLINLSHISGLIILAGLUSERLEUTHIRHLISLEUGIYLVSVALPROPOLAVALL	1867
Db	5821	CAGAAAGCATGAATTTGAATCTTTGTATACAAACGGGCAAGGGGCCCTCGCTGTATT	5880
QY	1868	IIEPROPOLAALAPROLEUSERGIYARGARGARGAPROTHIRLYSERSELYGISER	1887
Db	5881	ATTCGCCAGCGCTCCCTTCAGGGAGAAAGCAGACCCACTTAAAGCAAGGACG	5940
QY	1888	LYSSERSEARGSESERSELEUGIYANULYSERPROGILUENBERGIYASMLEUSER	1907
Db	5941	AAATTTACTTCAGAGACTTCTTGCGGAATAAAGCCCCCGACTTTCAGGTACCTGTCT	6000
QY	1908	GLYGINSERIALAASERVALLEUHIAPROGINLINHIREUHI.SPROPROGIYASMLE	1927
Db	6001	GGTCAGAGTCAAGCTTCACTGTTCGACCCCGACGAGACCCCTCCACCTCTCGGCAATC	6060
QY	1928	PROGIUSERGIYGINASNGINLEULEUGINPROLEULYSPOSERPROSESERASPAAN	1947
Db	6061	COAGAGTCGGGCGAGAAATCACTGTTCACGCCCTTAAAGCCATCTCCCTCAAGTGAAC	6120
QY	1948	LEUTHYSERIALAPHETHIRSESPGIYALALLESERVALPROSELEUSERALAPROGIY	1967
Db	6121	CTCTATTACGCTTCCACAGTAGTGTGCCATTTCAGTACCAAGCCTTTCGCTCCAGCT	6180
QY	1968	GLINGIYHRSERSETHIRASPTHVALGIYALATHYVALANSERGINALALGINALA	1987
Db	6181	CAAGGACCCAGACACACAAACACTGTGGGGCAACACTGAAACGCCCAAGCCGCAAGCT	6240
QY	1988	GLINPROPOLAMECTHIRSESERSEARGIYGLIYTHIRPETHIRASPALEUHI.SLYBLEU	2007
Db	6241	CAGCTCCTCGCAAGAGCTCCAGACGAGAAAGGACCATTCACAGTGACTTGCACAAAGTTG	6300
QY	2008	VALASPAANTPRALARGASPVALAMECASMLEUSERGIYARGARGIYSELYSGIYHS	2027
Db	6301	GTAACAAATTGGGCCGAGATGCCATGAATCTCTCAAGCAGAGAGAAACAAAGGGCAC	6360
QY	2028	MEFASNTYRIGUGIYPROGIYMETALARGULYBHESESRALAPROGIYGINLEUCYSILE	2047
Db	6361	ATGATATTATAGGGCCCTCGAAATGGCAAGAGATCTCTTGCACTGGGCAACGTGCATC	6420
QY	2048	SETHIRSESRANLEUGIYGIYSERIALAPROILESERIALALASERIALATHIRSELEU	2067
Db	6421	TCCATGACCTCGAACCTGGGAGGCTCTGCCCCCATCTCTGACGACATAGCTTACTCTTA	6480
QY	2068	GLIYHISPHETHIRLYSERMETCYSPIROPROGINLINYGLYPHEPROALATHIRPROPE	2087

Db	6461	GGTACTCTTACACCAAGTCATATGTCGCCGCCCCCAACGAGTATGCGTTTCCAGCTTACCCCATTT	65400
QY	2088	GLyAlaGlnTyrSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlnPheGlnPro	2107
Db	6541	GCGCGTCAATGGAGTGGGACGGGTGGGCCACGACACACAGCCACTGTGGCCAGTTCCAACT	6600
QY	2108	ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnIleSerIleSerAsn	2127
Db	6601	GtGGGAACtTGcCTcCTTcGcAGAAATTTcCAAcTcGAGcATTtTGcAGAAATTCATcAGCAAC	6660
QY	2128	ProProGlySerAsnLeuArgThrThr	2136
Db	6661	CCCCAGGCTCCAACTCGCGSACCACT	6667
RESULT 4	AR302588		
LOCUS	AR302588	7149 bp	DNA
DEFINITION	Sequence 1 from patent US 6541252.		linear
ACCESSION	AR302588		
VERSION	AR302588.1	GI:3169086	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 7149)		
TITLE	Walke, D.W., Hilpun, E., Donoho, G. and Turner, C.A. Jr.		
JOURNAL	Human kinases and polynucleotides encoding the same		
FEATURES	Patent: US 6541252-A 1 01-APR-2003;		
source	Location/Qualifiers		
	1..7149		

```

ORIGIN
/organism="unknown"
/mol_type="genomic DNA"

Alignment Scores:
Pred. No.:      1,17e-254      Length:      7149
Score:          10677.00        Matches:      2136
Percent Similarity: 89.67%      Conservative: 0
Best Local Similarity: 89.67%      Mismatch:     0
Query Match:     98.75%      Indels:      246
BB:              6            Gaps:         1

```

US-10-010-720-14 (1-2136) x AR302588 (1-7149)

QY 1 MetSerGlyValAlaIalaglutylglnserSerrThrProglYserLeuPheLeuSerPro 20
Db 1 ATGTCTGGGGGGCGCCGACAGAACGACGACACTCCCCGGTTCCCTTCTTCGCCG 60

QY 21 ProAlaProAlaProLYsaSnlgYserSerSerSaPsarSerValGlYglutylseuglY 40
Db 61 CCGGCTCTTGCCCCCAAGAATGGCTCCAGCTCCGATTCTCCTGGGGGAAMAATGGGA 120

QY 41 AlaAlaAlaAlaAspaAlaValThrGlYargThrglngluYrYArgArgArgHisThr 60
Db 121 GCCGCggcgccgcgcgcgtGTgAcCGcGcAgAcCGAGAGtACAgsccgcgcgcCACct 180

QY 61 MetAspLYsaPsarSerArglYAlaAlaAlaThrThrThrThrglnHIsaYrPhe 80
Db 181 ATGACAAAGACAGACGGCTGTGGCGCGCGCCGACCACTAACCACACTGAGACCGCTTCTTC 240

QY 81 ArgArgServallleCYsaPsarSerSnlaThrAlaleugluleuProglYleuProleu 100
Db 241 CGCGGAGAGGTATCTGCATCTCCATCCCATCTGACTGAGAGCTTCCGGGCTTCTCTT 300

QY 101 SerLeuProglInProSerSerleProAlaAlaValProglInserralProProgluProHis 120
Db 301 TCCCTGCCCCCAcCCAGACATCCCcCGcGTCTCCCGcAGAGTGTCCAACCGAGccccCAC 360

QY 121 ArgGlugluThrValThrAlaThrAlaThrserGlnValAlaglnglnInProproAlaAla 140
Db 361 CGGAGAGAACcGTAAcCGCACCGcCACTTCCcAGGTAGcCGACAGcCTTCCAGCGcCT 420

QY	141	AlAlaAProGluGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr	160
Db	421	GCGCCCTCTGGGAAACGGCCGTGCG3GGCCCTGCCCCCTGCATGTCCTCCAGACGATCC	480
QY	161	SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProPro	180
Db	481	AGCAAAAGCCGCCAGTGTCCCACTAGCTTGTGGGAGCAAAAGAGGAGCGCGCCG	540
QY	181	AlaArgSerGlySerGlyGlySerAlaLysGluProGlnGluArgSerGlnGln	200
Db	541	GCGAAGATGGCAACGGCGCGCGACGGCCCAAGAGCAACAGGAGGAACGAGCCACAG	600
QY	201	GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe	220
Db	601	CAGATGATATCGAAGAGCTGGAAACCAAGCCCTGGGAAATGTATACATAGAGCGCTTT	660
QY	221	LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp	240
Db	661	CTCAAGTTTGAATCGAAATCGGCAAGAGCTCTTTAAGACGGCTTCAAAAGCTGTGAC	720
QY	241	ThrGluThrThrValGluValAlaTTPCySGluLeuGlnAspArgLysLeuThrLysSer	260
Db	721	ACTGAACACACCGGGAAGTCGCTGGGTGTACATGACAGATCGAAATTAACAAAGTCT	780
QY	261	GluArgGlnArgPheLysGluGlnGluMetLeuLysGlyLeuGlnHisProAsnIle	280
Db	781	GAGAGGCAAGATTTAAAGAAAGCTGAATGTTAAAGGTCTTCAGCATCCCAATATT	840
QY	281	ValArgPheTyrAspSerTTPGlyLysSerThrValLysGlyLysCysIleValLeuVal	300
Db	841	GTTAGATTTTATGATTCCTGGGATCCACATTAAGGAAGAAGTGTGATTGTTTGATG	900
QY	301	ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys	320
Db	901	ACTGAACTTATGACGTCTGGAACCTTAAACGATATCGAAAGGTTTAAAGTATGAAG	960
QY	321	IleLysValLeuLysSerTTPCyArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr	340
Db	961	ATCAAAAGTTCTAAAGAGCTGGTGGCTGACATCTTAAAGGCTTCAGTTCTTCATACT	1020
QY	341	ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro	360
Db	1021	CGAATCCACTATCATTCACCGCGATTTAAATGTGACAAATCTTTATCACCGGCCCT	1080
QY	361	ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla	380
Db	1081	ACTGGCTCAGTCAGATGTGAGAACTCGGTCTGGCAACCTCGAAGCGGCTTCTTTGGCC	1140
QY	381	LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp	400
Db	1141	AAGATGTGATAGGATACCCACAGATTCATGGCCCTCGAAGTGTAGAGAAATATGAT	1200
QY	401	GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr	420
Db	1201	GAATCCGTGACGTTATGCTTTTGGATGTGCATGCTTTAGATGGCTACATCTGAATAT	1260
QY	421	ProTyrSerGluCysGlnAsnAlaIleGlnIleTyrArgArgValThrSerGlyValLys	440
Db	1261	CCTTACTCGAAGTGCCAAAATGCTGCACAGATCTACCGTCCGGTGAACAGGGGGTGAAG	1320
QY	441	ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle	460
Db	1321	CCAGCAGTTTGAACAAAGTACAAATTCCTGAAGTGAAGAAATATTGAAGAGATGACTA	1380
QY	461	ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln	480
Db	1381	CGACAAACAAAGATGAAGATATTCATCAAAAGCCTTTTGACCAATGCTCTTCCAA	1440
QY	481	GluGluThrGlyValArgValGluLeuAlaGluLysIleAspAspGlyGlyLysIleAlaIle	500
Db	1441	GAGGAAACAGGAGTACGGGTGTAATTTAGCAAGAAAGATGATGAGAAATAATATACCTTA	1500
QY	501	LysLeuTTPLeuArgIleGluAspIleLysLysLysLysGlyLysTyrLysAspAsnGlu	520

Db	1501	AAATTATGGCTAGCTATGAGTATATTAGAAATTAAAGGAAAAATACAAAGTAAATGAA	1560
Qy	521	AlAlIeGlupheserPheaspleuGluaAgAspValProGlupAspValaGInGInIuWec	540
Db	1561	GCATTTCAGTTTCCTTCTTTGATTATTAAGAGAGAGATGTCCTCCAGAAATGTTGCAAGAAATG	1620
Qy	541	ValGlusercIyTyValCysGInGlyAspHisAlTyThMerLAlaYsAlaIleTyAsp	560
Db	1621	GTAGAGCTGGGATATGTCCTGTGAAAGTGATCAAGACCATGGCTAAAGTATCAAAAGAC	1680
Qy	561	ArgValSerdeuIleTySArgIySArgIuGInArgGInIeuValArgGInGInGlu	580
Db	1681	AGAGTATCATTAATTATTAAGAGAAACGAGACGACGCGAGTTGGTACGGGAGAGCAAGAA	1740
Qy	581	LysIySArgIuGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn	600
Db	1741	AAAAAAGAGAGAGAGACCACTTCMAACACAGGTAGAACAAATCCAGTCTCCACAG	1800
Qy	601	ThrgIyIleTySglIneuProSerAlaSerThrgIyIleProThraIaSerThrThser	620
Db	1801	ACAGGATTCAGACAGCTCCCTTGTGCTAGACCGGACATACCTACTGCTTCACTTCA	1860
Qy	621	AlaSerValSerThrgInValGluProGInGInProGInIuIaaspGInHisGInGInIeu	640
Db	1861	GCTTCAGTTTCACACAGTAGAAACCTGAAAGAACCTGAGGACAGATCAACATCAACA	1920
Qy	641	GInTyrgInGInProSerIleSerValIeuSerAspGlyThryValaAspSerGlyGInGly	660
Db	1921	CAGTACCAAGAACCCAGATATATCTGCTGTATCTGAAGGACGGTTGACAGTGGTCAGGA	1980
Qy	661	SeSerSerValPheThrgInuSerArgValSerSerGInGInThryValSerTyrgIySerGIn	680
Db	1981	TCCCTCTCTTCCACAGAACTCCAGTAGAGCAGCAACAGACAGTTTCAATATGTTCCAA	2040
Qy	681	HisGInGInAlaHisSerThrgIyThryValProGlyHisIleProSerThryValGInAla	700
Db	2041	CATMAACAGGACACATCTTACAGGACAGCTCCAGGGCATATACCTTCTACTGTCCAGCA	2100
Qy	701	GInSerGInProHisGlyValTyProProSerSerValAlaGInGlyGInSerGInGly	720
Db	2101	CAGTCTAGCCCATGGGGGATATCCACCTCAAGTGTGGCAGAGGGGACAGACCGAGGT	2160
Qy	721	GInProSerSerSerSerIeuThrgIyValSerSerSerGInProIleGInHisProGIn	740
Db	2161	CAGCATCTCTCAATACCTTAAACGGGGTTTCATCTTCCCAACCATCAACATCTCCAG	2220
Qy	741	GInGInGInGlyIleGInGInThryAlaProProGInGInThryValGInTySerIeuSer	760
Db	2221	CAGCAGCAGGAAATACAGCAGACAGCCCTCTCCMAACAGAGCAGATTCACCTTCA	2280
Qy	761	GInThySerThrySerSerGInAlaThThThAlaGInProValSerGInProGInAlaPro	780
Db	2281	CAGACATCAACCTCAAGTAGGACCATCTGCAACGCAAGTGAATCAGCTCAAGCTCCA	2340
Qy	781	GInValIeuProGInValSerAlaGlyIySglIn-----	791
Db	2341	CAAGTCTTGCTCAAGTATCAGCTGAGAAACAGCTTCAGTTTCCAGACGATACCAACT	2400
Qy	791	-----	791
Db	2401	ATCCAAAGCGAAGCTCAGATCCAGTTGCGACACAACTTCGGTTGTTCAGTCCACTCT	2460
Qy	791	-----	791
Db	2461	GGTGTCTATTCTTCCAGTGGGACAGCGCTCCCTACTCCCTTGTCTCCTCAGTACCT	2520
Qy	791	-----	791
Db	2521	GTCTCTCAGATTCCATATCAACTCTCATGTGTCTACGGCTCAGACAGGTTTCTCATCC	2580
Qy	791	-----	791

Db 2581 CTTCCCATCATGATGGCAGCTGGCATTTACTGACGCTCTGCTCAGCTTGCTTCACTCTGCT 2640
 QY 791 ----- 791
 Db 2641 ACAACAGCTGCATGCCGGGGGTATCAACTGTTCTTCTAGTCACTTCCAACTCTCTG 2700
 QY 791 ----- 791
 Db 2701 CAGCTGTGACTCAGCTGCACAGTCAGTTCCACCAAGCTCTCAACACAGCAGTTCAG 2760
 QY 791 ----- 791
 Db 2761 TCCATGGGAATTCACAGCTAACCTTGGACAGCTGCTGAGTTCCACTTCTCTGGAGAT 2820
 QY 791 ----- 791
 Db 2821 GTTCTGTACAGGGCTTCCACCTGACAGTCCACACAGTACCCAGAGATTCAATATT 2880
 QY 791 ----- 791
 Db 2881 GCTCCCTCTCCACAGTGGCTTCTGTTGCACTTCTACAGTCTTACCCCTCCATG 2940
 QY 791 ----- 791
 Db 2941 CCGACAGAAGTACCTGCTACACCTGGGTACTTTCCACAGTGTGACCTTATGTGAA 3000
 QY 791 ----- 791
 Db 3001 TCAATCTTTAGTCTCTATGGTGTGTAGAGACAGCTTCAAGTCTCCACGACGA 3060
 QY 792 ----- SerThrgln 794
 Db 3061 GGGAGTTTAGCAACAGCCCACTACATCTCCAGCAAGCAGTTTGGAGTACTCAG 3120
 QY 795 G1YAlSerGlnValAlaProAlaGlnProValAlaValAlaGlnProGlnAlaThrgln 814
 Db 3121 GGAATCTCTCAGGTTGCTCTGACAGCAGCTGTGAGTGAAGCAAGCCCAAGCTTACCCG 3180
 QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 Db 3181 CCGACCACTTGGCTCTCTGTAGACAGTGCACATTCAGATGTTGCTTCAGGTATGAT 3240
 QY 835 AspGlyAsnGlnAsnValProSerSerSerGlyArgHisGlyArgThrThrLysArg 854
 Db 3241 GATGGCAATGAGAACGTCCTCATCTCCAGTGAAGGCAATGAGAGAACTTCAAAAACGG 3300
 QY 855 HisThrArgLysSerValArgSerArgSerArgHisGlyLysThrSerArgProLysLeu 874
 Db 3301 CATTAACGAAATCTGTAGAGAGTGTCTGACATGAAATACTTCAAGCCCAAAATTA 3360
 QY 875 ArgTlleLeuAsnValSerAsnLysGlyAspArgValAlaGlyCysGlnLeuGluThris 894
 Db 3361 AGAATTTTGAATGTTTCAATTAAGAGAGACCGAGTAGTGAATTCATTTAGAACTCAT 3420
 QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGlyGluTlleAla 914
 Db 3421 AATGAGAAATGATTCAATTCATTAATTCAGTATGATGTCACCCCGAGAGATGACA 3480
 QY 915 ThrIleMetValAsnAspPheTlleLeuAlaTlleGluArgGlySerPheValAspGln 934
 Db 3481 ACAATTTGTGAAACATGACTTTATTTCTGCAATTAAGAGAGAGGTGTTGTGATCA 3540
 QY 935 ValArgGluTlleTlleGlyLysAlaAspGlyMetLeuSerGluAspValSerValGluPro 954
 Db 3541 GTGGGAAATTTATGAAAAAGCTGATGAATGCTCAGTACGATGTCAGTGTGAAACA 3600
 QY 955 GluGlyAspGlnGlyLeuGlyLysSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
 Db 3601 GAGGTGATCATGAGGATTGAGAGAGTCAAGAGAAAGATGACTATGCTTTTCAGTTCT 3660
 QY 975 GlnLysLeuGlnGlyLysPheLysGlnProTlleProAlaSerSerMetProGlnGlnTlle 994
 Db 3661 CAGAAATTGAGAGAGAGCTTCAACCAACCAATTCCTGCTTCCATGCCACAGCAATA 3720

QY 995 GlyTlleProThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgArgPheTlleVal 1014
 Db 3721 GGCAATCTTACCACTTCTTTACTCAAGTTGTTCATTTCTGGGGAAGGGGTTTATAGTG 3780
 QY 1015 SerProValProGlySerArgLeuArgGlySerLysValPheProSerGluTlleThraP 1034
 Db 3781 AGTCTGTGCAAGAAAGCCGATTTAGAGAAATCAAAAGTTTCCAGTGAATTAACAGAT 3840
 QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
 Db 3841 ACAGTTGCTGCTCTACAGCTCAAGCCCTGAAATGAATCTGTCTCACTTGCATCATCC 3900
 QY 1055 LeuSerLeuGlnGlnAlaPheSerGlyLeuArgArgAlaGlnMetThrGlyLysProAsn 1074
 Db 3901 CTTAGTCTACAAAGCCCTTTTCTGAACTTGAAGTGTCCAAATGACAAAGAGACCCAA 3960
 QY 1075 ThrAlaProAsnAspPheSerHisThrGlyProThrPheProValAlaProPhePheLeu 1094
 Db 3961 ACAGACCTCCAACTTTAGTCATACAGAGCAACATTTCCAGTAGTACCTCTTCTTA 4020
 QY 1095 SerSerTlleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
 Db 4021 AGTAGCAATGCTGAGAGTCCCAACCAAGCAGCAGCAGCAGCAGCAGCTCTGCAACAGC 4080
 QY 1115 SerProProAsnAspLysSerThrSerValTlleGlnSerGlyValThrValProThrgln 1134
 Db 4081 AGCCCTCTTAATGACATTTCCATCTCATGATTAAGTGTGAGTTTACAGTCCCACTGAA 4140
 QY 1135 GluGlyTlleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLysProTlle 1154
 Db 4141 GAGGGAGATTGCTGAGTTGCCACACAGCAGCTGTGTAATTTCAAGTGTCTCCCACTA 4200
 QY 1155 ProProValSerGlySerProValLeuSerSerValAlaSerSerTlleThrTlleProAla 1174
 Db 4201 CCACCTGTGCTGATACACAGTACTTCCAGCTTCAAGTTTCAAGTATACAAATCTGCA 4260
 QY 1175 ValValSerTlleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluTlle 1194
 Db 4261 GTTGTCTCATATTTACTCATCTCCCGTCACTTCAAGTCCCAATCCACATCTGAGATC 4320
 QY 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAlaSerAla 1214
 Db 4321 GTTGTCTCTAGTACAGACCTGATCTTCAAGTACAGTTCAGCACTTCAAGCTCTGCA 4380
 QY 1215 GlyGlySerThrAlaThrProGlyProLysProProAlaValAlaSerGlnGlnAlaAla 1234
 Db 4381 GGGGGCAGTACTGCTACCCAGAGTCTTCAAGCTTCAAGTATCTCAGCAGGCGACA 4440
 QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
 Db 4441 GGGAGCACTACTGTGGAGGCCATTTACATCACTGATTTCTTCCACCACTTCAATCCAGC 4500
 QY 1255 ThrAlaSerGlnLeuSerTlleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlu 1274
 Db 4501 ACAGCTTCAACAGCTGTCACTTCAAGTTCAGAGAGTACTTCACTCTTACTTAAAGTAA 4560
 QY 1275 ThrValValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu 1294
 Db 4561 ACCGTGTATGTTCGACACACTCACTAGTAAGACATCTCATAGCAGTACCAACTGAGATTG 4620
 QY 1295 AlAspSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
 Db 4621 GCTTCTCTCTCTCTGACCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4680
 QY 1315 TlleSerGlnProGlyGlyLeuHisProLeuValTlleProSerValTlleAlaSerThrPro 1334
 Db 4681 ATTCTCAGCCTGTGGGCTGACATCTTGTGATTCATTCAGAGATAGCTTACTCTCT 4740
 QY 1335 TlleLeuProGlnAlaAlaGlyProThrThrSerThrProLeuLeuProGlnValProSerTlle 1354
 Db 4741 ATTCTTCCCAACAGAGAGACCTACTTCTTACACCTTATTAACCCCAAGTACTAGTATC 4800

QY 1355 ProProleuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuLeuHisSer 1374
Db 4801 CCACCCCTGGACAGCCCTGGCCCAAGTGGCTGTGACAGCAGCACTAATTCATAGT 4860
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer 1394
Db 4861 CAGCCTCAACAGCCTTCTCCCAACAGCCCACTCACTCATTTGCTGAGTAGATTC 4920
QY 1395 AsnThrGlnProValAlaProGlyValLeuAspIleLeuThrLeuGlnGlnLeuArg 1414
Db 4921 GATACACACCCCAAGCTCTCGAAATTGATGACATTAAGACTCTAGAAGAAAGCTGGG 4980
QY 1415 SerLeuPheSerGlnHisSerSerSerGlyValGlnHisAlaSerValSerLeuGlnThr 1434
Db 4881 TCTCTGTTCACTGAACACAGCTCATCTGAGCTCAGCATGCTCTGTCTCACTGGAAGCC 5040
QY 1435 SerLeuValIleGlnSerThrValThrProGlyValLeuProThrAlaAlaProSer 1454
Db 5041 TCACCTAGCATAGAGAGCAGCTGTCAACAGGACATCCCACTAGCTGTTGACCAAGC 5100
QY 1455 LysLeuLeuThrSerThrThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 5101 AAACCTCCGACTTCTACCAAGATCTTGCTTACCAACCACTTAACCACTAGAACCA 5160
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 5161 GTTGCTTTGCCAGTTACACAGTGTGTACACTGGGCAAGTTTCAACCCAGTCAAGCAT 5220
QY 1495 ThrThrSerGlyValValProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 5221 ACTACATCAGAGGTAAACCTGGAACTGCTCCCTCCAGACCACTCTAATTAAGGCTCG 5280
QY 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 5281 GTGCGCAGAGGGGACTGAACTTCCACAGGTACTTCAACCGAGCAGCTGCACT 5340
QY 1535 PheProGlyProSerSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAlaGlnLeu 1554
Db 5341 TTTCCAGAGCCTTCTCTAACCCAGTCCAGCACTCTAGAGATCTTGATGCTCAATTG 5400
QY 1555 ArgArgThrLeuSerProGluValIleThrValThrSerAlaValGlyProValSerMet 1574
Db 5401 AGAAGAACACTTACGTCAGAGATKATKATCAAGTGACTTCCGCGTGTGCTGTCCATG 5460
QY 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 5461 GCGGCTCAACAGCAGATACAGAACAGAGAACAGCCTCAGAAAGGTTCTCAAGTC 5520
QY 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 5521 AAAGAGGCCCTGTCTCTAGCACTAGTTCAGAGAGCTGTGTTTAAAGATGGGAGCATTT 5580
QY 1615 GlnValSerValAlaAlaAspGlyValGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
Db 5581 CAGGTTTCTGTTCAGCAGCAGCAGTCCAGAAAGAGGTTAAATAAGTCAGAAATATCA 5640
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerValLeuSerSerSerPro 1654
Db 5641 AAGTGTTCATTGTAATTCAGCAGCCTCAGAGTCTCAGTCTCATCAGTAGTAGTCA 5700
QY 1655 GlnSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyValLeuSerAsp 1674
Db 5701 GAGATTAACCTTGTAACCAAGCCGGAATGGCATTAACCATCCCTGTATCTCTTCAGAT 5760
QY 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnPro 1694
Db 5761 GTGCCAGAGAGTGCACAAACTACTGCTCAGAGGCAAAAGTCAGACACTGGGAGCCT 5820
QY 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 5821 ACCAAGGTTGACGTTTTCAGGTTGACACTTACAGCAAAAGAGGAGGTTTCTCTGTA 5880
QY 1715 SerLysThrGlnAspLysIleThrAspThrLysLysGlnGlyProValAlaSerProPro 1734

Db 5881 TCATAAACTAGAGCAAGATACAGACAAAGAAAGAGGACAGTGGCATCTCTCTCT 5940
QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlnLysProGlu 1754
Db 5941 TTTATGATTTGGAACAGGCTTCTCTGCTGTGATACCAAGAAAGAAAGAACCTGAA 6000
QY 1755 LeuSerGlnProSerHisLysAsnGlyProSerSerAspProGluAlaAlaPheLeuSer 1774
Db 6001 CTGTCAAGCCTTCACTTAATGAGCGCTCTCTGACCCGAGGCGCTTTTAAATG 6060
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerLysSer 1794
Db 6061 AGGATGAGATGATGTTGTCGAGTATGTCACACTGCCCCATCAGCTGAGCTCAAGAGC 6120
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerLysAsnSerPheAsnSerSerLysMetSer 1814
Db 6121 CTTCCTAGCCAGAAATCTAAGTCAAGCCTTAGTATTAATTAATTAATTAATTAATTA 6180
QY 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuArgArgLeuArg 1834
Db 6181 AGCGACATAGTACAGATATGAAAGATGACATTAAGTTAGAGCTGCGACGACTAGCA 6240
QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
Db 6241 GATAAATCATTCAAAGAGATTCAAGACCTGCAGAGTCCGAGAGCATGAATTAATCT 6300
QY 1855 LeuLysThrLysLeuGlyLysValProProAlaValIleLeuProProAlaAlaProLeu 1874
Db 6301 TTGTATACCAAACTGGGGAAGTGCCTCTGCTTATTAATTAATTAATTAATTAATTA 6360
QY 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
Db 6361 TCAGGAGAACACACACAGACCTCAAAAGCAAGGACAAATCTAGCGAAGCATTTCC 6420
QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
Db 6421 TTGGGAAATAAAGGCCCTCCAGCTTCAAGTAACTGTGTGTCAGAGTCAAGCTTCAATC 6480
QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
Db 6481 TTGACCCCCAGCAGACCTTCACCTCTCTGCAACATCCAGAGTCCGGGCGAGATAG 6540
QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
Db 6541 CTGTTACAGCCCTTAAGCATCTCCCTCCAGTGAACACTTATTAAGCTTCAACAGT 6600
QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
Db 6601 GATGGTGCATTTCACTACGACCAAGCTTCTCTCAGAGTCAAGGAAACGACAGCAAAAC 6660
QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnIleGlnProProAlaMetThrSer 1994
Db 6661 ACTGTTGGGGCAACGTAAGAGCCCAAGCCCAAGTCCAGCTCTGCGCATGACGCTC 6720
QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAsp 2014
Db 6721 AGCAGGAAGGACATTCACAGATGACTTCACAAAGTTGTAAGACATTTGGGCCGAGAT 6780
QY 2015 AlaMetAsnLeuSerGlyArgArgLysSerLysGlyHisMetAsnTrpGlnGlyProGly 2034
Db 6781 GCCATGAATCTCTCAGCAGCAGAGAGAAAGCAAAAGGCGACATGAATTAAGAGGCCCTGGA 6840
QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
Db 6841 ATGGCAAGAAAGTTCTGCACTTGAGGCACTGTCACTCATCATCACTCGAAGCTGGAT 6900
QY 2055 GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074
Db 6901 GGCTCTGCCCATCTCTGCGAGCATCAGTACTCTTAGAGCATTTACCAAGCTTAGT 6960
QY 2075 CysProProGlnGlnLysGlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThr 2094

/number=16
 3664. .3721
 /gene="WNK1"
 /number=17
 3722. .3844
 /gene="WNK1"
 /number=18
 3845. .5280
 /gene="WNK1"
 /number=19
 5281. .5364
 /gene="WNK1"
 /number=20
 5365. .5448
 /gene="WNK1"
 /number=21
 5449. .5509
 /gene="WNK1"
 /number=22
 5510. .5583
 /gene="WNK1"
 /number=23
 5584. .6245
 /gene="WNK1"
 /number=24
 6246. .6448
 /gene="WNK1"
 /number=25
 6449. .6643
 /gene="WNK1"
 /number=26
 6644. .6831
 /gene="WNK1"
 /number=27
 6832. .7149
 /gene="WNK1"
 /number=28

ORIGIN

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

Alignment Scores:

Pred. No.: 1.17e-254
 Score: 10677.00
 Percent Similarity: 89.63%
 Best Local Similarity: 89.63%
 Query Match: 98.75%
 DB: 9
 Gaps: 1

US-10-010-720-14 (1-2136) X HSA296290 (1-7149)

QY 1 MetSerGIyGlyAlaAlaGluIyGlnSerThrProGlySerLeuPheIeuSerPro 20
 Db 1 ATGTCTGGGGCGCGCAGAGAGACAGACACCTCCGGTTCCTGTTCTCTGCGCG 60
 QY 21 ProAlaProAlaProIyAsnGlySerSerSerSerSerValGlyGluIySleuGly 40
 Db 61 CCGGCTCTGCCCCCAAGATGGCTCCAGCTCCGATTCCTCCGTGGGGAGAAACTGGGA 120
 QY 41 AlaAlaAlaAlaAspAlaValThrGlyArgThrGluGlyTyrArgArgArgHisThr 60
 Db 121 GCCGGCGCGCGCAGCGCTGTGACCGCGCAGGACCGAGGTACAGCGCGCGCCACACT 180
 QY 61 MetAspIyAspSerSerArgIyAlaAlaAlaIaThrThrThrThrThrGluHisArgPhePhe 80
 Db 181 ATGGACAAAGACAGCGGTGGGGCGCGCGACCACTACCACTGAGGACCGCTTCTTC 240
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProIeu 100
 Db 241 CCGCGGAGGCTGATCTGGAGCTCCCAATGCACTGGCGTGGAGCTCCCGGCTTCTCTT 300
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
 Db 301 TCCCTGCCCCAGCCAGCATCCCGCGGCTGTCCGCAAGTGTCTCCACCGGAGCCCGCAC 360

QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGluValAlaGlnGlnProProAlaAla 140
 Db 361 CCGGAGAGACCGTGAACCGCCACCGCACTTCCAGGTAGCCAGCGCTCCAGCGCT 420
 QY 141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 421 GCCGCCCTGGGGAACAGCGCGCTGCGGGCCCTGCCCTCCCTGACGTGCCAGAGTACC 480
 QY 161 SerIyAspArgProValSerGlnProSerIeuValGlySerIySgluGluProProPro 180
 Db 481 AGCAAAAGCCGCCAGGTGCCAGCTTACCTTGTGGGAGCAAAAGAGAGACCGCGCG 540
 QY 181 AlaArgSerGlySerGlyGlySerAlaIySgluProGlnGluArgSerGlnGln 200
 Db 541 GCGAAGATGGACGCGCGCGCGGACCGCAAGAGCCACAGAGAGAGAGAGAGAGAGAG 600
 QY 201 GluAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
 Db 601 CAGATGATATCGAAGAGCTGAGACCAAGCGCTGGGGAATGTCTAACATGAGCGCTTT 660
 QY 221 LeuIySphAspIleGluIleGlyArgGlySerPheIyThrValTyrLysGlyLeuAsp 240
 Db 661 CTCAAGTTTACATCGAAATGCGAGAGCTCTTTAAGCGGTCTACAAAGGTGTGAC 720
 QY 241 ThrGluThrThrValGluValAlaIaTrpCysGluLeuGlnAspArgIySleuThrLysSer 260
 Db 721 ACTGAACCAACCGTGGAAAGTGGCTGTGTGAATCGACAGATTCGAAATTAACAAAGCT 780
 QY 261 GluArgGlnArgPheIySgluGluAlaGluMetLeuIySgluLeuGlnHisProAsnIle 280
 Db 781 GAGAGCGCAGATTTAAAGAAAGAGCTGAATGTAAAGGTCTTCACACATCCCAATTT 840
 QY 281 ValArgPheIyAspSerThrGlnSerThrValIySgluIySgluValIleuVal 300
 Db 841 GTTAGATTATATGATCTCTGGGAATCCACAGATTAAGAAAGAGATGATCTTTGGTG 900
 QY 301 ThrGluLeuMetThrSerGlyThrLeuIySgluThrIyLeuIySgluArgPheValMetLys 320
 Db 901 ACTGAAGTTATGAGCGTGTGGAACACTTAACGATCTGAAGAGTTTAAAGTATGATGAG 960
 QY 321 IleIySValLeuAspSerTrpCysArgGlnIleLeuIySgluLeuGlnPheLeuHisThr 340
 Db 961 ATCAAGTTTCAAGAGAGCTGTGGCTGCGTCAAGTCTTAAGGTCTTCAATCTTCAACT 1020
 QY 341 ArgThrProProIleIleHisArgAspLeuIySgluCysAspAsnIlePheIleThrGlyPro 360
 Db 1021 CGAAGTCCAGCTATCATTCACCGCGATCTTAATGTGACACATCTTATACCGGCGCT 1080
 QY 361 ThrGlySerValIySgluIyAspLeuGlyLeuAlaThrLeuIySgluAspSerPheAla 380
 Db 1081 ACTGGCTCAGTCAAGATTTGGAGACTCGGTCTGGCAACCTGAAACCGGCGCTTCTTGGCC 1140
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluIySgluYrAsp 400
 Db 1141 AAGAGTGTGATAGTAAACCCAGAGTTCATGGCCCTGAGATGATAGAGAGAAATATGAT 1200
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluIyYr 420
 Db 1201 GATTCGTTGAGCTTATAGCTCTTGGAGTGTGACGCTTGAGATGGCTACATCTGAATAT 1260
 QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValIySglu 440
 Db 1261 CTTACTCGAGAGTGCCAAAATCTCGAGATCTACCTCGCGGTGACAGTGGGTGAAG 1320
 QY 441 ProAlaSerPheAspIyValAlaIleProGluValIySgluIleIleGluGlyCysIle 460
 Db 1321 CCAGCCAGTTTGAACAAATGACCAATCTCTGAAGAGAGAAATTAATGAAGAGTGCATA 1380
 QY 461 ArgGlnAsnIyAspGluArgTyrSerIleIySgluAspLeuAsnIleAlaPhePheGln 480
 Db 1381 CGACAAACCAAGATGAAGATATTCATCAAAAGCCTTTGACACCAAGCTTCTTCCAA 1440
 QY 481 GluGluThrThrValArgValGluLeuAlaGluGluAspAspGlyGluIySgluIleAlaIle 500

Db	1441	GAGGAACAGAGATGCGGTAGAAATTGACAGAGAAATGATGAGAAATAATACCCATA	1500
QY	501	LysLeuTrpLeuArgIleGluAspIleLysIleLysGlyLysTyrLysAspAsnGlu	520
Db	1501	AAATTATGGCTACGATATTGAATATTAAAGAAATTAAGGAAATTAACAAAGATTAATGA	1560
QY	521	AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet	540
Db	1561	GCTATTGAGTTTCTTTTGTATTAGAGAGAGATGTCACAGAAAGATGTTGACAAAGAAATG	1620
QY	541	ValGluSerGlyTyrValCysGluGluLysAspHisLysIleMetAlaLysAlaLysAsp	560
Db	1621	GTAAGCTGGGATATGCTGTGAAGGTATCAACAGCATGGCTAAAGCTATCAAAAGAC	1680
QY	561	ArgValSerLeuIleLysArgLysArgLysGluGlnArgGlnLeuValArgGlnGluGln	580
Db	1681	AGAGATCATTAATTAAGAGAAACAGAGACGCGCATGGTGTGTAAGGAGAGCAAGAA	1740
QY	581	LysIleLysGlnGluGluSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln	600
Db	1741	AAAAAAGCAGAGAGAGAGAGAGCTCTCAACAGCAGGTAGAAACAATCCAGTCTCCAG	1800
QY	601	ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer	620
Db	1801	ACAGGAATCAAGCAAGCTCTCTCTGTCTAGACCGGCATACCTACTGCTTTCACCACTCA	1860
QY	621	AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu	640
Db	1861	GCTTCAGTTTCTACACAGATGAGAACCTGACAGACCTGAGGACATCAACATCAACACTA	1920
QY	641	GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyTyrValAspSerGlyGlnGly	660
Db	1921	CAGTACCAAGCAACCCAGATATCTGTGTATCTGATGGGACGTTGACAGTGTGACAGGA	1980
QY	661	SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln	680
Db	1981	TCCTCTGCTTCAACAGATCTCGAGTGGAGCCAGCAACAGACAGATTCATATGATGCCAA	2040
QY	681	HisGlnGlnAlaHisSerThrGlyTyrValProGlyHisIleProSerThrValGlnAla	700
Db	2041	CATGAACAGGCACATTTCTACAGGACAGCATGCCAGGACATATCTTCTCTGTCACAGCA	2100
QY	701	GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly	720
Db	2101	CAGCTCAGCCCATGGGGTATATCCACCTCAAGTGTGGACAGGGGACAGCCAGGGT	2160
QY	721	GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740
Db	2161	CAGCATCTCTCAAGTACCTTAACAGGGGCTTTCATCTTCCCAACCCATACAAACATCCCTCAG	2220
QY	741	GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer	760
Db	2221	CAGCAGCAGGGAATACAGCAGACAGCCCTCTCTCAACAGACAGTACGATTTTCACTTTCA	2280
QY	761	GlnHisSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	2281	CAGCACAACAACCTCCAGTGGAGGCCACTACTGACACAGCCAGTGAAGTCAAGCTCAAGCTCA	2340
QY	781	GlnValLeuProGlnValSerAlaGlyLysGln-----	791
Db	2341	CAAGTCTTGCTCAAGTATCAAGCTGGAAACAGCTTCCAGTTCCAGCAGTACCAACT	2400
QY	791	-----	791
Db	2401	ATCCAAGGCAACCTCAGATCCAGTTGCGACACAAACCTCGGTGTTCAGTCCACTCT	2460
QY	791	-----	791
Db	2461	GGTCTCATTTCTTCCAGTGGAGACAGCCGCTCTACTCTCTGTGCTCCTCAGTACCT	2520
QY	791	-----	791
Db	2521	GTCTCTCAGATTCCCATATCAACTCTCATGTGTCTACGGCTCAGACAGGTTTCTCATCC	2580
QY	791	-----	791
Db	2581	CTTCCCATCAAAATGGCAGCTGGCATTAAGCTCAGCTCTGTCTCAGCTGGCTTCACTGTCT	2640
QY	791	-----	791
Db	2641	ACAAAGCTGGATCCCGGGGATACAACTGTGTCTTCTAGTACGTTCCAACTCTTCTG	2700
QY	791	-----	791
Db	2701	CAGCCTGACTCAGCTGCCAAGTCAGGTTCACCCACAGCTCTCAACACAGCACTTTCAG	2760
QY	791	-----	791
Db	2761	TTCATGGGAATACAGCTAACCTTGGACAACTGTGAGTTCCACTTCTCTGAGAT	2820
QY	791	-----	791
Db	2821	GTTCTGTACAGGCTTCCCACTTGACTGCCACCAAGTACCCAGAGATTCAAATATT	2880
QY	791	-----	791
Db	2881	GCTCCCTTTCACAGTGTGCTTGTGTGATTCATTCAGTCCATCCCTCCAGT	2940
QY	791	-----	791
Db	2941	CCGACAGAAATGACTGGCTACACTGGGTACTTCCACAGTGTGACGCTTATGTGAA	3000
QY	791	-----	791
Db	3001	TCAATCTTTTACTTCTCATGTGTGTGTAGAGAGACAGTTCAAGTGTCCAGCCAGGA	3060
QY	792	-----	792
Db	3061	GAGAGTTTACACAGACCCCACTACCTCTCCAGACAGATTTGGAGATGACTCAG	3120
QY	795	GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln	814
Db	3121	GAGTCTCTGAGTGTCTCTGACAGACAGTGTGACATACACAGCCCAAGCTTACACAG	818
QY	815	ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer	834
Db	3181	CCGACCACTTGGCTTCTCTGTAGACAGTGCACATTCATGCTTCAAGTATGAGT	838
QY	835	AspGlyAsnGlnAsnValProSerSerSerGlyArgHisGlnGlyArgThrThrLysArg	854
Db	3241	GATGGCAATGAGAACGTCCTTCCAGTGGAAAGCATGAAAGAAAGAACCTTACAAAACGG	858
QY	855	HisTyrArgLysSerValArgSerArgSerArgHisGlnLysThrSerArgProLysLeu	874
Db	3301	CATTACCGAAATCTGTGAAGAGTCTCTGACATGAGAAACCTTCAACGCCCAAAATTA	878
QY	875	ArgIleLeuAsnValSerAsnLysGlyAspArgValValGluCysGlnLeuGlnThrHis	894
Db	3361	AGAAATTTGATATTTCAATTAAGAGACCGAGTGTAGAAATGCTCAATTAAGACCTCAT	900
QY	895	AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGlnIleAla	914
Db	3421	AATAGGAAATGGTTACATTCATTAATTTGACTGTAGTGTGACAAACCCGAGGATAGCA	918
QY	915	ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln	934
Db	3481	ACAAATATGATGAACATGACTTATTTATTAAGAAATGAGAGAGAGTGTGTGATCA	938
QY	935	ValArgGluIleIleGlnLysAlaAspGluMetLeuSerGluAspValSerValGluPro	954
Db	3541	GTGCGAGAAATTAATTAAGAGTGTGAAATGCTGATGAGATGTCTAGTGTGAGAACCA	958
QY	955	GluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer	974
Db	3601	GAGGGTATCAGGATGAGTGTACAAAGGAAAGATGACTATGCTTTCAGGTTCT	978

QY 975 GlnIysLeuGlnIysGlnPheIysGlnProIleProAlaSerSerMetProGlnGlnIle 994
Db 3661 CAGAAATTTGAGAGAGGTTCAACCAACCAATTCCTGCTTCACAGCCACAGCAATA 3720
QY 995 GlyIleProThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgPheIleVal 1014
Db 3721 GGCAATTCCTACAGATTCTTTACTCAAGTTGTTCTTCTGCGGGAAGCGTTTATAGTG 3780
QY 1015 SerProValProGlnSerArgLeuArgLeuSerIysValPheProSerGluIleThrAsp 1034
Db 3781 AGTCTGTCGAGCAAGAGCCGATTACGAAATCAAAAGTTTCCCAAGTGAATTAACGAT 3840
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetLeuLeuSerHisSerAlaSerSer 1054
Db 3841 ACAATTCCTGCTTACAGCTCAGAGCCGAGAAATGAATCTGTCTCACTGCACTATCC 3900
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGlnGlyProAsn 1074
Db 3901 CTTAGTCTACAAAGCCCTTTTCTGAACTTAGAGTGGCCAAATGACAGAAAGCCCAAC 3960
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProPheLeu 1094
Db 3961 ACAGACCTCCAACTTAGTCATACAGAGCAACCAATTCCAGTAGTACCTCTTCTTA 4020
QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaThrAlaProValProAlaThrSer 1114
Db 4021 AGTAGCATGTGTGAGTCCCAACACAGACAGCCAGCCAGCAAGTCCCTGCAACAGC 4080
QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 4081 AGCCCTCTTAATGACATTTCCACATCACTATTCAGTGTGAGGTAGAGTCCCACTGAA 4140
QY 1135 GlnGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLeuProIle 1154
Db 4141 GAGGGGATGTGCTGAGTGCACACAGCAGAGTGTGTACTTCCAGGTGCTCCCACTA 4200
QY 1155 ProProValSerGlnSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 4201 CCACCTGTGTGATCAACAGATCTTCCAGCCGATGATTCAGATACATACATCTGCA 4260
QY 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 4261 GTTGCTCATATCTACTACATCCCCGTCATTCAGTCAAGTCCCACTCACTGAGATC 4320
QY 1195 ValValSerSerThrAlaLeuThrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 4321 GTTGTTCTAGTACAGCACTGATCTTCAGTACAGATTTCAGCAACTTCAGCCTCTGCA 4380
QY 1215 GlyGlySerThrAlaThrProGlyProIysProProAlaValValSerGlnGlnAlaIle 1234
Db 4381 GGGGGCAGTACTGTAACCCAGAGTCTTAAGCCTCAAGCTGATGATCTCAGAGCAGCA 4440
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrPheProSer 1254
Db 4441 GGACACACTACTGCTGGAGCCCATTAACATCAGTTTCTACCAACCATTCATTCACAGC 4500
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrLeuAlaGlu 1274
Db 4501 ACAGCTTCAACAGCTGCTCATTGAGCTTACAGACAGTACTTCACTCACTTACTGAA 4560
QY 1275 ThrValValValSerAlaHisSerLeuAspIysThrSerHisSerSerThrThrGlyLeu 1294
Db 4561 ACCGGTAGTAGTACGACACACTCAGTACGATTAAGACATTCATACAGCATCACTGATG 4620
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerThr 1314
Db 4621 GCTTTCTCCCTCTCTGCAACATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4680
QY 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4681 ATTCTCAGCTGTGGCTGATCTTTGTCATTCATCAGTATGATACCTTCTACTCTCT 4740
QY 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 4741 ATTCTTCCCAAGCAGAGAGACTTCTTACACTTATTTATCCCAAGTACAGTATC 4800
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4801 CCACCTTGTGTACAGCTGTGGCCAAATGTCTCTCTTACAGCAGACATCAATTCATAGT 4860
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValAspSer 1394
Db 4861 CAGCTTCAACGAGCTTGTCTTCCCAACAGCCCATCTCATTTGCTCTGAGTATGATTTCT 4920
QY 1395 AspThrGlnProIysAlaProGlyIleAspAspIleIysThrLeuGlnGlnIysLeuArg 1414
Db 4921 GATACACACCCAAAGCTCTGGAAATTGATGACATAAAGCTTAGAAGAAAGAACTGGGG 4980
QY 1415 SerLeuPheSerGlnHisSerSerSerGlyValAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 4981 TCTCTGTTCAGTGAACAGCTCATCTGAGCTCAGATGCTCTGTCTCTCTCTCTCTCTCT 5040
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 5041 TCACTAGCATGAGAGACACTGTACACAGCAGATCCCACTACTGCTGTGACACAGC 5100
QY 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 5101 AAATCTGACTTACCAACAGTACTTGTACCAACCAACATTAACCTAGAGAA 5160
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 5161 GTTGCTTGGCAGTATACACAGAGTGTACACCTGGGCAAGTTTCTACCCAGTACGACT 5220
QY 1495 ThrThrSerGlyValIysProGlyThrAlaProSerIysProProLeuThrIysAlaPro 1514
Db 5221 ACTCATCATGAGAGTGAACCTGGAACCTCTCTCCCAAGCAGCTCTTACTAAGGCTCG 5280
QY 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnLeuProPro 1534
Db 5281 GTGCTGCAGTGGGTAAGTCTTCCAGAGTACTTCAACAGAGGAGGAGCAGCTGCACT 5340
QY 1535 PheProGlyProSerLeuThrGlnSerGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 5341 TTTCAGGACCTTCTCAACCCAGTCCAGCACTCTTGAAGGATCTTGAAGTCAATTG 5400
QY 1555 ArgArgThrLeuSerProGlu**IleThrValThrSerAlaValGlyProValSerMet 1574
Db 5401 AGAGAAACATTAAGTCAAGATTAATCAAGTACTTGGCGTTGCTGTGCTGTCANTG 5460
QY 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnIysGlyValSerGlnVal 1594
Db 5461 GCGGCTCCAAAGCAATCAGAGACAGAGACAGAGCCTCAGAAAGGAGTTCCTCAAGTGC 5520
QY 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheIysMetGlyArgPhe 1614
Db 5521 AAAGAAAGCCCTGCTCTTCAAGTCAATGTTCAAGAGTGTGTTTAAAGTGAAGATT 5580
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnIysGlnGlyIysAsnIysSerGlnAspAla 1634
Db 5581 CAGGTTTCTGTGTCAGCAGAGCGGTGCCAGAAAGGGGTAATAAATATCAGAAATGCA 5640
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerSerValLeuSerSerSerPro 1654
Db 5641 AAGTCTGTTCATTTGAATCCAGCAGCTCAGAGTCTTCAAGTCTATCAAGTAGTATCCA 5700
QY 1655 GlnSerThrLeuValIysProGlnIysProGlnIysIleThrIleProGlyIleSerSerAsp 1674
Db 5701 GAGAGTACTCTGTGTAACCAAGCCGAAATGCAATACATCTCTGTATCTTCAAGT 5760
QY 1675 ValProGlnSerAlaHisIysThrThrAlaSerGlnAlaIysSerAspThrGlyGlnPro 1694
Db 5761 GTGCCAGAGAGTCCCAAAATCTGCTCTAGAGGCAAGATCAGACACTGGGAGCCT 5820
QY 1695 ThrIysValGlyArgPheGlnValThrThrThrAlaAsnIysValGlyArgPheSerVal 1714

Db 5821 ACCAAGTTGGACGTTTTCAGTGAACACTACAGCAACAAAGTGGCTGTTCTCTGTA 5880
 Qy 1715 SerLysThrGluAspLysIleThrAspThrLysLysGluGlyProValAlaSerPro 1734
 Db 5881 TCAAAACAGGAGACAGATCACTGACACAAAGAAAGAACAGCTGACCTCTCTCT 5940
 Qy 1735 PheMetAspLeuGluGlnAlaValLeuProAlaValIleProLysLysGluLysProGlu 1754
 Db 5941 TTATAGATTGGAAACAGCTGTTCTCTGCTGATGATCAAGAAAGAAAGAGCTGTA 6000
 Qy 1755 LeuSerGluProSerHisIleLeuasnGlyProSerSerAspProGluAlaAlaPheLeuSer 1774
 Db 6001 CTGTACAGGCTTTCACATCTAAATGAGCGCTTCTTGACCCCGAGCCGCTTTTAAATG 6060
 Qy 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
 Db 6061 AGGATGTGATGATGATGTTCCGGTAGTCTCACACTGCCCCATCAGCTGAGCTCAAGAGC 6120
 Qy 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSer 1814
 Db 6121 CTTCCTAGCCCAAGATCTAAGTCAAGCTTAACTAGTAACTTAACTCCTCTTACATAGT 6180
 Qy 1815 SerAspAsnGluSerAspIleGluAspGluAspLeuLysLysGluLeuArgArgLeuArg 1834
 Db 6181 AGCGCAATGAGTCAATGATATGCAAGATGAAGATTAAAGTTAAGACTGCGACGACTACGA 6240
 Qy 1835 AspLysHisLeuLysGluIleGlnAspLeuGlnSerAspArgGlnLysHisGluIleGlnSer 1854
 Db 6241 GATAAACAATCTCAAAAGAGATTCAGAGACTGACAGAGTCCCCAGAGAGATGAATTTGAATCT 6300
 Qy 1855 LeuThrThrLysLysGluLysValProProAlaValIleIleProProAlaAlaProLeu 1874
 Db 6301 TTGTATACCAAACTGGGCAAGTGGCCCTGCTGTTATTTATTTCCCGGCGTCTCCCTT 6360
 Qy 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
 Db 6361 TCGGGAGAGAGACAGACGACCCCACTAAAGCAAAAGCCAAATCTAGTCGAGACAGATTCC 6420
 Qy 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
 Db 6421 TTGGGGAATATAAAGCCCCACCTTTCAGTAACTGTTGATTCAGAGTGCAGCTTCCAGTTC 6480
 Qy 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlySerGlyGlnAsnGln 1934
 Db 6481 TTGCACCCCCGACGAGACCTCCACCTCTGCAAACTCCAGAGTCCGGGCAAGATCAG 6540
 Qy 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
 Db 6541 CTGTTACAGCCCCCTTAAGCATCTCCCTCCAGTACAACTTATTCAGACTTCCACAGT 6600
 Qy 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
 Db 6601 GATGTGTCATTTCAGTACCAAGCTTTCCTCTCCAGGTCAAGAGAACAGAGACAAAC 6660
 Qy 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnIleGlnProProAlaMetThrSer 1994
 Db 6661 ACTGTTGGGGCAACAGTGAACAGCCCAAGCCGCCAAGCTCAGCTCCGCAATGAGCTCC 6720
 Qy 1995 SerArgLysGlyThrPheThrAspAspLeuHisIleLysLeuValAspAsnThrPalaArgAsp 2014
 Db 6721 AGCAGGAAGGGCAATTCACAGATGACTTGACAAATGGTGAACAATTTGGGCCCGAGAT 6780
 Qy 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnLysGlnGlyProGly 2034
 Db 6781 GCCATGATCTCTCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6840
 Qy 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
 Db 6841 ATGCAAGAGAAATCTCTGCACTGGGCAACTGTGATCTCCAGTCCGAGACCTGGGT 6900
 Qy 2055 GlySerLalaProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074

Db 6901 GGCCTGCCCCCATCTCTGACGATCAGCTACTCTAGTCACTTCACCAAGTCTAG 6960
 Qy 2075 CysProProGlnGlnIntyGlyPheProAlaThrProPheGlyAlaGlnIntySerGlyThr 2094
 Db 6961 TGCCCCCAAGCAGATAGGCTTTCACACTCACTCACTCACTTGGGCTCAATGAGTGGAGC 7020
 Qy 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
 Db 7021 GGTGGCCAGACACAGAGCCACTTGGCCAGTTCACACTGTGGAGACTGCTCTTGGCAG 7080
 Qy 2115 AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg 2134
 Db 7081 AATTTCACATCAGCAATTTGAGAAATCCATCAGCAACCCCGAGCTCCAACTGGCG 7140
 Qy 2135 ThrThr 2136
 Db 7141 ACCACT 7146
 RESULT 6
 AR302602 6327 bp DNA linear PAT 12-JUN-2003
 LOCUS AR302602
 DEFINITION Sequence 29 from patent US 6541252.
 ACCESSION AR302602
 VERSION AR302602.1 GI:31690900
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 REFERENCE
 1 (bases 1 to 6327)
 AUTHORS Walke,D.W., Hilpun,E., Donoho,G. and Turner,C.A. Jr.
 TITLE Human kinases and polynucleotides encoding the same
 JOURNAL Patent: US 6541252-A 29 01-APR-2003;
 FEATURES
 source
 1..6327
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,976-254 Length: 6327
 Score: 10646.00 Matches: 2108
 Percent Similarity: 98.69% Conservative: 0
 Best Local Similarity: 98.69% Mismatches: 0
 Query Match: 98.46% Indels: 28
 DB: 6 Gaps: 1
 US-10-010-720-14 (1-2136) x AR302602 (1-6327)
 Qy 1 MetSerGlyGlyAlaAlaGluLysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 1 ATGTCTGGCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
 Qy 21 ProAlaProAlaProLysAsnGlySerSerSerSerAspSerSerValGlyLysLysGly 40
 Db 61 CCGGCTCTGCCCCCAAGATGAGCTCCAGCTCCGATTCCTCGTGGGGAGAAACTGGGA 120
 Qy 41 AlaAlaAlaAlaAspAlaValIleThrGlyArgThrGluGlyIleArgArgArgArgIleThr 60
 Db 121 GCGGCGCGCGCGAGAGTGTGACCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 Qy 61 MetAspLysAspSerArgGlyAlaAlaAlaAlaThrThrThrThrThrThrGluHisArgPhePhe 80
 Db 181 ATGAGCAAGAGACAGCCGTGGGGGGCGCGGACCACTACCACTGAGCACTGAGCTTCTTC 240
 Qy 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 Db 241 CGCCGAGCGCTCATCTGCGACTCCAAATGCACTGAGCACTGAGCTTCCGCGCTCTCT 300
 Qy 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
 Db 301 TCCCTGCCCCAGCCAGCATCCCGGCGCTGTCCGACAGAGTCTCCACCGAGAGCCCAAC 360
 Qy 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140

QY 861 ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
 Db 2497 AGAGATCGCTCTCGACATGAAAAAACTTCACGCCCAAAATTAAGATTGTAATGTTTCA 2556
 QY 881 AsnLysGlyAspArgValValGluCysGlnLeuGluThrHisAsnArgLysMetValThr 900
 Db 2557 AATAAAGGACCGAGTAGTAAGTAATGTAATGAGACTCAATTAAGAAAAATGTGTTCA 2616
 QY 901 PheLysPheAspLeuAspGlyAspAsnProGlnGluIleAlaThrIleMetValAsnAsn 920
 Db 2617 TTCAATTTGACCTGATGATGTCACACCCGAGGAGATAGCAACAATTAATGTGTAACAT 2676
 QY 921 AspPheIleLeuAlaIleGluArgGlnSerPheValAspGlnValArgGluIleIleGlu 940
 Db 2677 GACTTATTTCTAGCAATAGAGAGAGAGATCGTTTGTGATCAAGTCGAGAAATTAATGAA 2736
 QY 941 LysAlaAspGlnMetLeuSerGlnAspValSerValGluProGlnGluLysAspGlnLysLeu 960
 Db 2737 AAAGCTGATGAAATGCTCAGTAGAGATGTCAGTGTGGAACCAAGGGTGATCAGAGGATTTG 2796
 QY 961 GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGlnGlyGlu 980
 Db 2797 GAGAGTCAACAGAGAAAGATGATGCTTTCAGTTCTCAGAAATTTGAAAGAGAG 2856
 QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
 Db 2857 TTCAAAACAACCAATTCCTCGCTCTTCATGCCACAGCAAAATAGGCAATTCCTACACGTTCT 2916
 QY 1001 LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGlnSer 1020
 Db 2917 TTAACCTAAGTTGTTCATTCGCGGGAAGCGGTTTATAGTAGAGTCCGTGCCAGAAAGC 2976
 QY 1021 ArgLeuArgGlnSerLysValPheProSerGlnIleThrAspThrValAlaAlaSerThr 1040
 Db 2977 CGATTACAGAAATCAAAAGTTTCCCGCAGTAATTAACAAATACAGTTGCTGCTCA 3036
 QY 1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnAla 1060
 Db 3037 GCTCAGAGCCCTGGATGAACTTGTCTCCTCATCTGCATCATCCCTTAGCTCAACAAGGCC 3096
 QY 1061 PheSerGlnLeuArgArgAlaGlnMetThrGlnGlyProAsnThrAlaProProAsnPhe 1080
 Db 3097 TTTCTGAACTTAAAGCTGCCAAATAGCAGAAAGAACCAAVACAGCACTCCCAAACTTT 3156
 QY 1081 SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGlyVal 1100
 Db 3157 AGTCAACAGAACCAACATTTCCAGTAGTACCTCTTCTTAAGTAGCATTTGCTGAGATC 3216
 QY 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
 Db 3217 CCAACCAAGCAGACGACACAGCAGCAGTCCCTGCACAAACAGCAGCCCTCAATAGACATT 3276
 QY 1121 SerHisSerValIleGlnSerGlnValThrValProThrGlnGlnGlyIleAlaGlyVal 1140
 Db 3277 TCCACATCAAGTAATTCAGTGTGAGGTTCACATGCCCCTGAAGAGGGATTTGCTGAGATT 3336
 QY 1141 AlaThrSerThrGlyValValThrSerGlyLysLeuProIleProProValSerGlnSer 1160
 Db 3337 GCCACCAAGCAGAGTGTGTAATTCAGTGTGTTCCCATACCACTGTGTCTGAATCA 3396
 QY 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
 Db 3397 CCAATACCTTCCACCGTAGTTTCAAGTATCACATACCTGCAAGTTGTCTCAATATCTACT 3456
 QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla 1200
 Db 3457 ACATCCCGCTACTTCAAGTCCCAACATCCACATCTAGATCGTTGTTTCTAGACAGCA 3516
 QY 1201 LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlyLysThrAlaThr 1220
 Db 3517 CTGATCTCTTCAGTAACAGTTTACGCAACTTCAGCCCTCTGAGGGGGCAGTACTGCTACC 3576

QY 1221 ProGlyProLysProProAlaValValSerGlnGlnAlaAlaGlySerThrThrValGly 1240
 Db 3577 CCAAGTCTTAAGCCTCCAGCTGTAGTATCTAGCAGGAGCAGCAGACATCTGTGGGA 3636
 QY 1241 AlaThrLeuThrSerSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
 Db 3637 GCCACATTAACATGATTTCTTACACACCACTTCATCCCAAGCAGCTTACAGCTGTCC 3696
 QY 1261 IleGlnLeuSerSerSerSerThrThrProThrLeuAlaGluThrValValValSerAla 1280
 Db 3697 ATTCAAGCTTACAGACGATCTTCTTACTTCACTTAAGTGAACCGTGTGTGAGCCCA 3756
 QY 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
 Db 3757 CACTCCTAGATTAAGACATCTCATAGCAGTACCACTGGATTTGGCTTTCTCTCTTCCA 3816
 QY 1301 ProSerSerSerSerSerProGlyValAlaGlyValSerSerTyrIleSerGlnProGlyLys 1320
 Db 3817 CCATCTTCTCTCTCTCTCTCGAGAGCAGAGTGTCTATATTCTCAGCCTGTGGG 3876
 QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
 Db 3877 CTGCATCTTGTGTCAATTCATCACTGATTAAGCTTCTTACTCTTATTTCTCCCAAGCAGCA 3936
 QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
 Db 3937 GGACCTACTTCTTACACCTTATTAATCCCAAGTACTGATACCACTTGTGTACAGCT 3996
 QY 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
 Db 3997 GTTCCCAATGCGCTCGTACAGACAGCACTAATTCATAGTCAAGCTCAACACGCTTTG 4056
 QY 1381 LeuProAsnGlnProHisThrHisCysProGlnValAlaAspSerAspThrGlnProLysAla 1400
 Db 4057 CTTCACCAAGCCCACTACTCATTTGCTCGAAGTAAATTCGTGATTAACAACAACCAAGCT 4116
 QY 1401 ProGlyIleAspAspIleLysThrLeuGlnGlnLysLeuArgSerLeuPheSerGlnHis 1420
 Db 4117 CTTGGAATTTGATGACATTAAGACATCTGAGAGAAAGCTGCGTCTCTGTTAGTGAACAC 4176
 QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
 Db 4177 AGCTCATCTGAGCTCAGCATGCTCTGTCTCACTGAGAGACTCACTAGTATAGAGAGC 4236
 QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
 Db 4237 ACTGTCAACAGGACATCCAAACTACGTGTGTGCACCAAGCAAACTCTGACTTCTAAC 4296
 QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
 Db 4297 ACAAGTACTTGTCTTACACCAACCAATTTACACTTGAACAGTGTGCTTCCAGTTACA 4356
 QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
 Db 4357 CCAAGTGTCAACCTGCGGCAAGTTTCAACCCAGTCAAGCACTACTACATCAGAGAGTAAAA 4416
 QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
 Db 4417 CTGGAAGCTGTCTCTCAAGCCACCACTTAAGCTCAAGGCTCGGGCTGCCAGTGGGATCT 4476
 QY 1521 GluLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu 1540
 Db 4477 GAACCTCCAGCAGGTACTTACCAAGGAGACAGTGCACCTTTTCCAGGACCTTCTCTA 4536
 QY 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgArgThrLeuSerPro 1560
 Db 4537 ACCCAAGTCCCAAGCACTCTAGAGGATCTGAAGCTCAATTTGGAAGAACACTTAGTCCA 4596
 QY 1561 Glu***IleThrValThrSerAlaValAlaGlyProValSerMetAlaAlaProThrAlaIle 1580
 Db 4597 GAGATATTCACAGTGAATCTTGCAGTTGCTGTGTGTGTCCATGCGGGCTCCAAAGCAGCATC 4656
 QY 1581 ThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGlnGlyProValLeu 1600

```

Db 4657 ACAGAGCAGAGAAACACAGCCTCAGAGGGTCTTCTCAGCTCAAGAGGAGGCGCTGTCTTA 4716
QY 1601 AAlaThrsSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db 4717 GCACTAGTTCAGAGAGCTGAGTGTCTTTTAAAGTGGGACGATTCAGAGTTCTGTGTGACGCA 4776
QY 1621 AspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAlaLysSerValHisPheGlu 1640
Db 4777 GACGGTGCCAGAGAAAGAGGTAATAATAGTCAGAAAGATGCAAGATCTGTTCATTGTGAA 4836
QY 1641 SerSerThrSerGlnSerSerValLysSerSerSerSerProGlnSerThrLysValLys 1660
Db 4837 TCCAGACACTCAGAGCTCCTCAGTGTCTATCAGTAGTACTCCAGAGAGACTTGGTGA 4896
QY 1661 ProGlnProAsnGlyLysThrLysProGlyLysSerSerAspValProGlnSerAlaHis 1680
Db 4897 CCAGAGCCGGAATGGATACCATCCCTGTATCTCTTCAGATGTGCCAGAGAGTGCCCA 4956
QY 1681 LysThrThrAlaSerGlnAlaLysSerAspThrGlyGlnProThrLysValGlyArgPhe 1700
Db 4957 AAACCTACTGCTCAGAGGCAAGTCAAGACACTGGGAGCCTACCAAGCTTGAGCTTTT 5016
QY 1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnAspLys 1720
Db 5017 CAGGTGACACTACGACCAACAAAGTGGTCTGTTCTGTATCAAAACAGAGGACAA 5076
QY 1721 LLeThrAspThrLysLysGlnGlyProValAlaSerProPheMetAspLeuGln 1740
Db 5077 ATCTCTGACAGAAAGAAAGAGAGACCACTGGCATCTCTCTTATGATTTGGACAA 5136
QY 1741 AlaValLeuProAlaValLysProLysGlnLysProGlnLysSerGlnProSerHis 1760
Db 5137 GCTGTTCTTCTCTGCTGTATACCAAGAAAGAGCCTGAACCTGTCAAGCCTTCAAT 5196
QY 1761 LeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db 5197 CTAAATGGGCGCTCTTGTGACCCGAGGCGCTTTTAAAGTAGGAGATGGATATGAT 5256
QY 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
Db 5257 TCCGGTAGTCCACACTCCGCCCATAGCTGAGCTCAAGAGGCTTCCACAGACATCTA 5316
QY 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSerSerAspAsnGlnSerAsp 1820
Db 5317 AGTCAAAACCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5376
QY 1821 LLeGlnAspGlnAspLeuLysLeuGlnLysLeuArgArgLeuArgAspLysHisLeuLysGln 1840
Db 5377 ATCGAAGATGAAGACTTAAGTAGTAGTGGACGACACTACGATTAACATCTCAAGAG 5436
QY 1841 LLeGlnAspLeuGlnSerArgGlnLysHisLeuLysLeuSerLeuThrLysLeuGln 1860
Db 5437 ATTCGAGACTTGCAAGATGCGCAGAGATGAATTTGATTTGATTAACCAAACTGGG 5496
QY 1861 LysValProProAlaValLysLeuProProAlaAlaProLeuSerGlyArgArgArgArg 1880
Db 5497 AAGGTGCCCTCGCTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5556
QY 1881 ProThrLysSerLysGlySerLysSerSerArgSerSerSerSerLeuGlyAsnLysSerPro 1900
Db 5557 CCCACTAAAGCAAGAGGAGCAAAATCTAGTCGAAACAGATCTCTGGGAAATTAAGCC 5616
QY 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
Db 5617 CAGCTTTCAAGTAACTGCTGTGGTCAAGATGCAAGTTCAGTTCACCCCGCAGACACC 5676
QY 1921 LeuHisProProGlyAlaAsnLysProGlnSerGlyGlnAsnGlnLeuLeuGlnProLeuLys 1940
Db 5677 CTCCACCTCTCTGGCAACATCCCAAGTCCGGGCAAGATCAGCTGTACAGCCCTTAAG 5736
QY 1941 ProSerProSerSerAspAsnLeuLysSerAlaPheThrSerAspGlyAlaAlaSerVal 1960

```

```

Db 5737 CCATCTCCCTCCAGTACCAACCTCTATTCAGCTTACCAAGATGATGATTCATTCAGTA 5796
QY 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db 5797 CCAAGCCTTCTGTCTCAGAGTCAAGGAAACAGAGCAACAACTGTGGGCAACAGTG 5856
QY 1981 AsnSerGlnAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
Db 5857 AACAGCCAGCCGCCCAAGCTCAGCTCCTCCCTCAGCTCAGCTCAGAGAGGACATTC 5916
QY 2001 ThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGly 2020
Db 5917 ACAGATGACTTGCAAGATTTGATGACATTTGGGCGGAGATGGCATGATCTCTCAGGC 5976
QY 2021 ArgArgGlySerLysGlyHisMetAsnThrGlnGlyProGlyMetAlaArgLysPheSer 2040
Db 5977 AGGAGAGAGCAAGGAGGACATTAATTAATTAATTAATTAATTAATTAATTAATTA 6036
QY 2041 AlaProGlyGlnLeuCysLysSerMetThrSerAsnLeuGlySerAlaProLysSer 2060
Db 6037 GCACCTGGGCAACTGTGCATCTCATGACCTCGAACCTGGGTGGCTGCGCCCATCTCT 6096
QY 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnThr 2080
Db 6097 GCAGCATCAGCTACTCTCTCTGATGCTCCTCAGCAAGTCTATGTCCTCCCAACAGAT 6156
QY 2081 GlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThrGlyGlyProAlaProGln 2100
Db 6157 GGGCTTTCAGCTACCCCAATTTGGCTCATATGAGTGGAGCGGTGGCCCAACAGCACAG 6216
QY 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnLysSerAsn 2120
Db 6217 CCACCTGGCCAGTTCACAACTGTGGAACTGCTCTCTGAGAAATTCATCAGCAAT 6276
QY 2121 LeuGlnLysSerLysSerAsnProProGlySerAsnLeuAspThrThr 2136
Db 6277 TTGCAGAAATCCATCAGCAACCCCGAGGCTCCAACTCGGAGACACT 6324

RESULT 7
AR302599 6606 bp DNA linear PAT 12-JUN-2003
LOCUS DEFINITION Sequence 23 from patent US 6541252.
ACCESSION AR302599
VERSION AR302599.1 GI:31690897
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE 1 (bases 1 to 6606)
AUTHORS Walke, D.W., Hilbun, E., Donoho, G. and Turner, C.A. Jr.
TITLE Human kinases and polynucleotides encoding the same
JOURNAL Patent: US 6541252-A 23 01-Apr-2003;
FEATURES
source Location/Qualifiers
1..6606
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 1.54e-252 Length: 6606
Pred. No.: 10589.50 Matches: 2108
Score: 94.57% Conservative: 0
Percent Similarity: 94.57% Mismatches: 0
Best Local Similarity: 97.94% Indels: 121
Query Match: 97.94% Gaps: 2
DB: 6

US-10-010-720-14 (1-2136) x AR302599 (1-6606)
QY 1 MetSerGlyGlyAlaAlaGlnLysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
Db 1 ATGTCTGGCGGGGCGGCGAGAGAGAGACAGACACTCCCGGATTCCTCTCTCGCGG 60
QY 21 ProAlaProAlaProLysAsnGlySerSerSerAspSerSerValGlyGlnLysLeuGly 40

```

```
Db 61 CCGGCTCTCGCCCAAGAAATGGCTTCAGCTCCGATTCCTCGTGGGAGAACTCGGA 120
Qy 41 AAlaAlaAlaAspAlaValThrGlyArgThrGluGluTyrArgArgArgArgThr 60
Db 121 GCGCGGCGCGCGGAGCTGTGACCGGAGAGACCGAGAGATACAGCGCGCGCGCACT 180
Qy 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrThrGluHisArgPhe 80
Db 181 ATGACAAAGAGACAGCGGTGGGCGCGCGGACCACTTACCACTGAGCAACCGGCTTCTC 240
Qy 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
Db 241 CGCGGAGCGTCATCTGACCTCCATGCCACTGACAGAGCTTCCCGGCTTCTCTT 300
Qy 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGlnProHis 120
Db 301 TCCGCGCCCAAGCCAGATCCCGCGGCTGTCCGAGAGTGTCCACCGAGCCGAC 360
Qy 121 ArgGluGluThrValThrAlaThrAlaThrSerGluValAlaGlnGlnProProAlaAla 140
Db 361 CGGAGAGAGACCGGAGCCGACCGCACTTCCGAGTAGCCAGAGACCTCCAGCGCT 420
Qy 141 AlaAlaProGluGluGluAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Db 421 GCGCGCCCTGGGAGAACAGCGCGTGGGCGCGCTGCGCTGACCTGCCAGAGTACC 480
Qy 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProProPro 180
Db 481 AGCAAAAGACCGCCAGGTCTCCAGCTTACCTTGTGGGAGACCAAGAGAGACCGCGCG 540
Qy 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGlnGluArgSerGlnGln 200
Db 541 GCGAGAAAGTGCAGCGCGCGCGAGCGCAAGCAAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 201 GlnAspAspIleGluGluLeuGluThrIleValAlaValGlyMetSerAsnAspGlyArgPhe 220
Db 601 CAGAGATGATCGAAGAGCTGAGAGACCAAGGCGGTGGAAATGCTTAAAGATGCGGCTT 660
Qy 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
Db 661 CTCAAGTTTGACATCGAAATCGAGAGAGCTCTTTAAGACGAGCTCAAAAGGCTGGAC 720
Qy 241 ThrGluThrThrValGluValAlaIleProCysGluLeuGlnAspArgLysLeuThrLysSer 260
Db 721 ACTAAACCAACCGGAGAGTGGCTGTGTGAATCGACGATCGAAAATTAAACAAGTCT 780
Qy 261 GlnArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280
Db 781 GAGAGGCGAGAGATTTAAAGAGAGAGTGAATGTAAAGGTCTTCAGCATCCCAATATT 840
Qy 281 ValArgPheLysAspSerIleProGluSerThrValLysGlyLysCysIleValLeuVal 300
Db 841 GTTAGATTTTATGATTCCTCGGAATCCACAGTAAAGAAAGAAAGATGATTTTGTG 900
Qy 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Db 901 ACTGAACCTTATGAGTGTGGAACCTTAAACGATCTGAAAAGGTTTAAAGTAGTAG 960
Qy 321 IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
Db 961 ATCAAACTTCTAAAGACTGTGCGCTGCAATCTTAAAGTCTTCAGTTTCTTCATACT 1020
Qy 341 ArgThrProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 1021 CGAAGCTCACCTATCATTCACCGGAGCTTAAAGTAGTACAAACTTTATCACCGGCGCT 1080
Qy 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
Db 1081 ACTGGGCTCAAGTCAAGATTGAGAGACTGCTGTGGCAACCTGGAAGCGGCTTCTTTGCC 1140
Qy 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
Db 1141 AAGAGTGTATAGTATACCCAGAGTTCATGGCCCCCTGATGATGAGAGAAATATGAT 1200
Qy 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db 1201 GAATTCGTTGACGTTTATGCTTTTGGAGTGGCATGCTTGAAGATGCTAATCTGAATAT 1260
Qy 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValAlaThrSerGlyValLys 440
Db 1261 CTTTACTCGAGTGCACAAATGCTGCACAGATCTACCGTCCGAGACAGAGGGGTGAAG 1320
Qy 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
Db 1321 CCAGCGAGTTTGAACAAGTAGCAATTCCTGAATGAAGAAATATTATGAAGATGATCA 1380
Qy 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
Db 1381 CGACAAACAAAGATGAAGAAATATTCATCAACAAACCTTTGAAACCAATGCTTCCAA 1440
Qy 481 GlnGluThrGlyValAlaArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
Db 1441 GAGGAAACAGAGATACCGGTATGATTCAGTACAGAAAGATGATGAGAAAAATAGCCATA 1500
Qy 501 LysLeuTyrLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
Db 1501 AAATTATGGCTACGTAATGAAGATATTAAAGAAATTAAGGAAATATCAAAAGATTATGA 1560
Qy 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 1561 GCTATTAGTCTTTCTTTGATTGAGAGAGATGTCACAGAAAGTGTGCAAGAAATG 1620
Qy 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 1621 GTAGAGTCTGGATATGCTGTGAAGATGATCACAAGACCATGCTTAAGTATCAAGAC 1680
Qy 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGluLeuValArgGluGluGln 580
Db 1681 AGAGTATCAATTAATTAAGAGAAACGAGACAGCGGAGTGTGTAAGGAGAGACCAAGAA 1740
Qy 581 LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln 600
Db 1741 AAAAAAAGCAGAGAGAGACAGTCTTAAACAGAGGTAAACAAATCAGTCTCCAG 1800
Qy 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1801 ACGAGATCAAGAGACTCCCTTCTGTACAGACCGGCACTACTACTCTTCCACTTCA 1860
Qy 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db 1861 GCTTCAGTTTCTACAAAGTAGAACTGAAAGAACTGAGGACATCAACATCAACAATA 1920
Qy 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1921 CAGTACCGAACCCAGATATATCTGTATCTGATGGGCGGTGACAGTGTGACAGGA 1980
Qy 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1981 TCCCTGTCTTCAAGATCTCAGATGAGAGAGCAACAGCAATTTATGTTGCTCCCA 2040
Qy 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2041 CATGAACAGGACATTTCAAGGACAGATCCAGGAGATATTAATCTTACTGTCAAGCA 2100
Qy 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 2101 CAGTTCAGCCCAATGGGATATATCCACCTCAAGTGTG----- 2139
Qy 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2139 ----- 2139
Qy 741 GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2140 ---CAGAGGGAATACAGAGAGAGAGCCCTCTCTCAACAGACAGTGAATTCATCTTCA 2196
```

QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2197 CAGACATCAACCTCCAGTAGAGGCCACTACGACAGCCAGAGTCAAGCTCCCA 2256
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 2257 CAAGCTTGGCTCAAGTATACGCTGGAAAACAGGGCTTCCACCTCCAGCTCCACACAG 2316
QY 791 ----- 791
Db 2317 TACCAGAGATCAATATATGCTCCCTCTTCCAAGTGGCTGTGTTCATCCATCT 2376
QY 791 ----- 791
Db 2377 ACAGTCTTACCCCTCCCATGCGACAGAGATCTGGCTACACCTGGTACTTCCACA 2436
QY 791 ----- 791
Db 2437 GTGGTGCAGCCTTATGTGGATCAATCTTTAGTTCCTATGGGTGGTGAAGAGACAG 2496
QY 791 ----- 791
Db 2497 GTTCAAGTGTCCAGCCAGAGAGGAGTTAGCACAGCCCCACTACATCTCCACGCA 2556
QY 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGlnProValAlaVal 807
Db 2557 GCAGTTTGGAGAGTACTCCAGAGAGTCTCTCAGGTGCTCTCGACAGCCAGCTTCAGTA 2616
QY 808 AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2617 GCACAGCCCAAGCTACACCAGCCAGCCACTTGGCTTCTCTGTGAGACAGTCAATTCA 2676
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGlnValProSerSerSerGlyValArgHis 847
Db 2677 GATGTGCTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2736
QY 848 GlnGlyValGlnThrThrValArgHisGlyValArgSerValArgSerValArgHis 867
Db 2737 GAAAGAAACATCAAAAGGAGATTAACGAAATATGTAAGAGAGTGGCTTCGACATGAA 2796
QY 868 LysThrSerArgProLysLeuValArgLysLeuValSerAsnLysGlyAspArgValVal 887
Db 2797 AAAATTTACCCGCCAAATTAAGATTTTGAATGTTTCAANAAAGAGACCGGATAGTA 2856
QY 888 GlnCysGlnLeuGlnThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly 907
Db 2857 GAATGTCAATTAGAGACATCATATAGGAAATGTTTACATTCAAATTTGACCTTACATGAT 2916
QY 908 AspAsnProGlnGlnLysLeuAlaThrLysMetValAsnAsnAspPheLysLeuAlaLysGln 927
Db 2917 GACAAACCCGAGAGATACCAACATTTATGTGACAAATGATCTTATTTCTTACAAATGAG 2976
QY 928 ArgGlnSerPheValAspGlnValArgLysLeuLysGlnLysValAspGlnLysLeuSer 947
Db 2977 AGAGAGTGGTTGTGATCAATGCGAGAAATTAATTAAGAAAAGCGATGAATGCTCAGT 3036
QY 948 GlnAspValSerValGlnProGlnGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAsp 967
Db 3037 GAGGATGTCAGTGTGAAACCAAGAGGTGATTCAGGAGTTCGAGAGCTTCACAGAAAGAT 3096
QY 968 AspTyrArgLysPheSerGlySerGlnLysLeuGlnGlyLysPheLysGlnProLysProLys 987
Db 3097 GACTATGGCTTTTACAGTCTTCAGAAATTTGAAAGAGAGTTCACAAACCAATTTCTCGG 3156
QY 988 SerSerMetProGlnGlnLysGlyLysProThrSerSerLeuThrGlnValValHisSer 1007
Db 3157 TTTTCCATGCCACAGAAATAGGATCTTCTACAGATTTTAACTCAAGTGTTCATTTCT 3216
QY 1008 AlaGlyValGlnPheLysValSerProValProGlnSerArgLeuArgLysSerVal 1027
Db 3217 GCGGGAAGGCGGTTTATGTGATGCTCTGCGCAAGAAACCGATTCAGAAATCAAAAGTT 3276

QY 1028 PheProSerGlnLysLeuThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn 1047
Db 3277 TTCCCGATGAATTAACAGATACGATTGCTGCTTACAGCTCAGAGCCCTGGAAATGAAAC 3336
QY 1048 LeuSerHisSerAlaSerSerLeuSerLeuGlnLysAlaPheSerGlnLysValArgVal 1067
Db 3337 TTGCTCAGCTGTGATCATCCCTTACGTCTAACAGAGGCTTTTCTGAACCTTACAGTGC 3396
QY 1068 GlnMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisGlnGlyProThrPhe 1087
Db 3397 CAATATGACAGAAAGACCAACAGACCTTCAAACTTATGATCAATACAGAGCAACATTT 3456
QY 1088 ProValValProProPheLeuSerSerLeuAlaGlyValProThrThrAlaAlaThr 1107
Db 3457 CCAATAGTACTCTTCTTTTATAGATATGCTGAGTCCCAACACAGAGAGCCACA 3516
QY 1108 AlaProValProAlaThrSerSerProProAsnAspLysSerThrSerValLysGlnSer 1127
Db 3517 GCACAGTCCCTGCACACAGAGCCCTCTAATGACATTTCCACATCAGTAAATCAGTCT 3576
QY 1128 GlnValThrValProThrGlnGlnGlyLysLeuAlaGlyValAlaThrSerThrGlyValVal 1147
Db 3577 GAGGTACAGTGCCTCACTGAAGAGGAGATGCTGAGTTCACACAGCAGGTGTGTA 3636
QY 1148 ThrSerGlyLysLeuProLysProProValSerGlnSerProValLeuSerSerValVal 1167
Db 3637 ACTTCAGGTGTCTCCCATACCACTGTGTCTGATCAGATCAGATCTTCCAGGCTAGTT 3696
QY 1168 SerSerLysThrLysProAlaValAlaSerLysSerThrThrSerProSerLeuGlnVal 1187
Db 3697 TCAAGTATCAAAATACGACAGTGTCTCAATATCTATCAATCCCTCACTTCAACTC 3756
QY 1188 ProThrSerThrSerGlnLysLeuValSerSerThrAlaLeuTyrProSerValThrVal 1207
Db 3757 CCAATTCACATCTGAGATGTTGTTTCTAGTACAGACGTATCTTCACTGATACAGTT 3816
QY 1208 SerAlaThrSerAlaSerAlaGlyLysSerThrAlaThrProGlyProLysProAla 1227
Db 3817 TCAGCAATTCAGCTCTGCAGAGGAGGAGTCTGATACCCAGAGTCTTACCTCCACT 3876
QY 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
Db 3877 GTAGATCTCAGAGGACGACGACGACACTGCTGGAGGACCAATTAACTCACTTCT 3936
QY 1248 ThrThrThrSerPheProSerThrAlaSerGlnLeuSerLysGlnLeuSerSerSerThr 1267
Db 3937 ACCACCACTTATTTCCAAAGCAGCTTCAACAGCTGTCACTTACCTTACGACGACT 3996
QY 1268 SerThrProThrLeuAlaGlnThrValValAlaSerAlaHisSerLeuAspLysThrSer 1287
Db 3997 TCTACTCTTACTTATAGGAAACCGGTGATGAGCGCACACTCATGATTAAGACATCT 4056
QY 1288 HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
Db 4057 CATAGCACTACAACTGATGCTGCTTCCCTCTCTGACCAATCTTCTCTCTCTCTCT 4116
QY 1308 GlnAlaGlyValSerSerTyrLysSerGlnProGlyGlyLysHisProLeuValLysPro 1327
Db 4117 GAGCAGAGAGTGTCTAGTATATTTCTACGCTGTGGCTGATCCTTTGGTCTATCCA 4176
QY 1328 SerValLysLeuAlaSerThrProLysProGlnAlaAlaGlyProThrSerThrProLeu 1347
Db 4177 TCAGTATAGCTTACTTACTTCTTCTTCCCAAGCAGAGACCTACTTACACCTTTA 4236
QY 1348 LeuProGlnValProSerLysProProLeuValGlnProValAlaAspValProAlaVal 1367
Db 4237 TTAACCCAGAGACCTAGATATCCACCTTGTGACGCTGTGTGCAAAATGCTGCTGTA 4296
QY 1368 GlnGlnThrLeuLysHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
Db 4297 CAGCAGACATTAATTCATATGATGAGCTTACACAGCTTGTCTTCCAAACGCCCAATCT 4356
QY 1388 HisCysProGlnValAspSerAspThrGlnProLysAlaProGlyLysAspAspLys 1407

D	4357	CATTGCTCGAAGTGAATTTCTATACACAAAGCTCTCGAATGATACATAAG	4416
Q	1408	ThreuleugluLysleuArgSerleuPheSerGluHisSerSerSerGlyValIleHis	1427
D	4417	ACTTAGAAGAAAGCTGGCTCTGTTCACTGAACACAGTCACTCTGGAGCTACAGT	4476
Q	1428	AlaSerValSerleuGluThrsSerleuValIleGluSerThrsValThrsProGlyLeuPro	1447
D	4477	GCTCTGCTCTCAGAGACCTCAGTACATAGAGAGACAGTCAACACAGCATCCCA	4536
Q	1448	ThrsThrAlaValAlaProSerLysleuThrsSerThrsThrsCysleuProPro	1467
D	4537	ACTACTGCTGTTGACACAGCAAACTCTGACTTCTACCAAGACTCTTACACCA	4596
Q	1468	ThrsleuLeuProleuGlyThrsValAlaLeuProValThrsProValThrsProGlyGln	1487
D	4597	ACCAATTTACACACTAGAGAAAGTCTCTTGCAGTTACACAGTGTCACTGGGCA	4656
Q	1488	ValSerThrsProValSerThrsThrsSerGlyValLysProGlyThrsAlaProSerLys	1507
D	4657	GTTTCTACCCAGTACAGCACTACTACATCAGGAGGAACTGGAACTGCTCCCTCAAG	4716
Q	1508	ProProleuThrsLysAlaProValleuProValGlyThrsGluLeuProAlaGlyThrsleu	1527
D	4717	CCACTCTAACCTAAAGGCTCCGGTGTGCTGAGGGGTACTGAACTTCCAGAGTACTCTA	4776
Q	1528	ProSerGluGlnLeuProProPheProGlyProSerleuThrsGlnSerGlnProleu	1547
D	4777	CCAGCGAGCAGCTGCCACTTTTCCAGAGCTTCTCTAACCCAGTCCCAAGCACTCTA	4836
Q	1548	GluAspLeuAspAlaGlnLeuArgArgThrsleuSerProGlu**IleThrsValThrsSer	1567
D	4837	GAGATCTTGTAGTGTCAATTAGAAGAACTTAGTCCAGATATATACAGTACTCTCT	4896
Q	1568	AlaValGlyProValSerMetAlaAlaProThrsAlaIleThrsGluAlaGlyThrsGlnPro	1587
D	4897	GCGGTGACCTGCTGCTCAGCTGCTCCACAGCAATCACAGCAAGCAAGCAAGCT	4956
Q	1588	GlnLysGlyLysSerGlnValLysGlnGlyProValleuAlaThrsSerSerGlyValAlaGly	1607
D	4957	CAGAAAGGCTGTTTCTCAAGTCAAGAAAGGCTCTGCTCAGCACTAGTCCAGAGCTGT	5016
Q	1608	ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyAlaGlnLysGlnGly	1627
D	5017	GTTTATAGATGGAGCGATTTCAAGTCTCTGTGACAGACAGCGGCCAGAAAGAGGT	5076
Q	1628	LysAsnLysSerGluAspAlaLysSerValHisPheGluSerSerThrsSerGluSerSer	1647
D	5077	AAAHATTAAGTCAGAAAGATGCAAAAGCTGTTCATTTGATCCACACACTCAGAGTCTCA	5136
Q	1648	ValleuSerSerSerSerProGluSerThrsleuValLysProGluProAsnGlyIleThrs	1667
D	5137	GTGCTATCAATAGTATAGTCCAGAGAGTACCTTGTGAAACCAAGAGCCGAATGGCAATACC	5196
Q	1668	IleProGlyLysSerSerAspValProGluSerAlaHisLysThrsThrsAlaSerGluAla	1687
D	5197	ATCCCTGGTATCTCTTCAAGATGTGCCAGAGAGTCCCAACAACTACGTGCTCAGAGCA	5256
Q	1688	LysSerAspThrsGlyGlnProThrsLysValGlyArgPheGlnValThrsThrsAlaAsn	1707
D	5257	AAAGTCAGACACTGGGCAAGCTTCAAGAGTGGAGGTTTTCAGGAGCAACACTACAGCAAC	5316
Q	1708	LysValGlyArgPheSerValSerLysThrsGluAspLysIleThrsAspThrsLysGlu	1727
D	5317	AAATGGTGGTCTTCTCTGATCAAAAACTAGAGCAAGATCACTGACCAAAAGAAAGAA	5376
Q	1728	GlyProValAlaSerProProPheMetAspLeuGlnAlaValleuProAlaValIle	1747
D	5377	GGACAGAGGCACTCTCTCTTTATGATTTGGAAACAAGCTTTCTCCGTGCTGATA	5436
Q	1748	ProLysLysGlnLysProGluLeuSerGluProSerHisleuAsnGlyProSerSerAsp	1767
D	5437	CCAAAGAAAGAGAGCTGAACTGTACAGAGCTTCACTATCAATAGGCGCTTCTGAC	5496
Q	1768	ProGluAlaAlaPheLeuSerArgAspValAlaAspGlySerGlySerProHisSerPro	1787
D	5497	CCGAGGCGCGCTTTTATAGTAGGAGATGGATGTGTGTTCCGTATGCCACACTCGGCC	5556
Q	1788	HisGlnLeuSerSerLysSerleuProSerGlnleuLeuSerGlnSerleuSerAsnSer	1807
D	5557	CATCAGCTAGACTCAAAAGAGCTTCTTAGCCAGAACTTAAGTCAAAGCCTTAGTAATTA	5616
Q	1808	PheAsnSerSerThrsMetSerSerAspAsnGluSerAspIleGluAspGluAspLeuLys	1827
D	5617	TTTAACTCTCTTATCATAGTAGGACACATAGATCATGATTCGAAGATGAAGACTTAAG	5676
Q	1828	LeuGluLeuArgArgLeuArgAspLysHisleuLysGluIleGlnAspLeuGlnSerArg	1847
D	5677	TTAGAGCTGCAGACATACGAGATTAACATCTCAAGAGATTCAGAGCTTCAGAGTGC	5736
Q	1848	GlnLysHisGluIleGlnSerleuThrsLysleuGlyLysValProProAlaValIle	1867
D	5737	CAGAAAGCATGAATTTGAATCTTTGTATACAACTGGGCAAGTCCCTGCTGTTATT	5796
Q	1868	IleProProAlaAlaProleuSerGlyArgArgArgArgProThrsSerLysGlySer	1887
D	5797	ATTCCTCCAGCTGCTCCCTTTCAGGAGAGAGACGACACCACTAAAGCAAGGACGC	5856
Q	1888	LysSerSerArgSerSerSerleuGlyAsnLysSerProGlnleuSerGlyAsnleuSer	1907
D	5857	AAATCTAGTCAGACACAGTCTCTTGGGGAATTAAGCCCCAGCTTTCAGGTACCTGTCT	5916
Q	1908	GlyGlnSerAlaAlaSerValleuHisProGlnGlnThrsleuHisProProGlyAsnIle	1927
D	5917	GGTACAGGTGAGCTTCACTTGTCCACCCAGAGACCCCTCCCTCTGGCAACATC	5976
Q	1928	ProGluSerGlyGlnAsnGlnleuLeuGlnProleuLysProSerProSerSerAspAsn	1947
D	5977	CCAGAGTCCGGGCAATACAGCTGTTCAGCCCTTAAGCCATCTCCCTCAGTACACAC	6036
Q	1948	LeuThrsSerAlaPheThrsSerAspGlyAlaIleSerValProSerleuSerAlaProGly	1967
D	6037	CTTATTCAGCTTACACAGATGTGTCATTTCAAGTACCAACCTTTCTGCTCCAGGT	6096
Q	1968	GlnGlyThrsSerSerThrsThrsThrsValGlyAlaIleThrsValAsnSerGlnAlaAla	1987
D	6097	CAAGCAACGAGCAGCAAAACACTGTTGGGCAACAATGAAACGCCCAAGCCCAAGCT	6156
Q	1988	GlnProProAlaMetThrsSerSerArgLysGlyThrsPheThrsAspLeuHisLysleu	2007
D	6157	CAGCTCTCCCATGACGTCCAGCAGAGAGGACATTCACAGATGACTTGCACAAATTG	6216
Q	2008	ValAspAsnThrsAlaArgAspAlaMetAsnLeuSerGlyArgArgGlySerLysGlyHis	2027
D	6217	GTACACATTTGGGCGCCAGATGCCATTAATCTTCAAGCAGAGAGGAAACCAAGGGCAC	6276
Q	2028	MetAsnLysGlnGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCysIle	2047
D	6277	ATGAATTAAGAGGCCCTTGAATGGCAAGAAAGTCTCTGCACTGGGCAACTGTGATC	6336
Q	2048	SerMetThrsSerAsnleuGlyLysSerAlaProIleSerAlaAlaSerAlaThrsSerleu	2067
D	6337	TCCATGACCTCGAACCTGGGTGGCTGCCCCCATCTCTCAGAGATCAGCTACCTCTCA	6396
Q	2068	GlyHisPheThrsLysSerMetCysProProGlnGlnThrsPheProAlaThrsProPhe	2087
D	6397	GGTCACTTCAACAAAGTATATGTGCCCCCAAGAGATATGCTTTCCAGCTACCCCAATT	6456
Q	2088	GlyAlaGlnThrsPheSerGlyThrsGlyProAlaProGlnProleuGlyGlnPheGlnPro	2107
D	6457	GGCGCTCAATGGAGTGGGAGCGGGGCCAGCAACAGGCACTTGGCCAGTTCACACT	6516
Q	2108	ValGlyThrsAlaSerleuGlnAsnPheAsnLysSerAsnleuGlnLysSerIleSerAsn	2127
D	6517	GTGGGAACCTGCTCTTGCAGAAATTCACATCAGCAATTGGCAAAATCCTACAGCAAC	6576

QY	561	ArgValSerLeu1LeysArgLysArgGlnAArgLnuLeuValArgLnuGlnGlnu	580
Db	1681	AGAGTATCTATTATTATTAAGAGAAACGAGAGCAGCGCAAGTTGGTACGGAGGACAGAA	1740
QY	581	LysLysLysGlnGlnGlnSerSerLeuLysGlnGlnValGlnGlnSerSerAlaSerGln	600
Db	1741	AAAAAAAAACGAGAAAGAGAGAGCTCTCAAAACAGAGGATGAAGCAATCCAGTCTTCCAG	1800
QY	601	ThrGly1LeLysGlnLeuProSerAlaSerThrGly1LeProThrAlaSerThrThrSer	620
Db	1801	ACAGGAATCAAGAGAGCTCCCTCTCTGTAGACACGGCATACCTACTGCTTCAACCACTTCA	1860
QY	621	AlaSerValSerThrGlnValGlnProGlnGlnProGlnAlaAspGlnHisGlnLeu	640
Db	1861	GCTTCAGTTTCAACACAGTAGAACCTCGAAGAACCTGAGGAGCATCAATCAACATCA	1920
QY	641	GlnTyrglnGlnProSer1LeSerValLeuSerAspGlyThrValAspSerGlyGlnGly	660
Db	1921	CAGTACCAACCAACCAATATCTGTATCTGATATCTGAATGGACGGTGAAGAGTGAAGGGA	1980
QY	661	SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrgLysGln	680
Db	1981	TCCTCTGTCTTCAAGAAATCTCGAGTGAAGACCAACAGACAGTTTCAATATGTTCCAA	2040
QY	681	HisGlnGlnAlaHisSerThrGlyThrValProGlyHisLleProSerThrValGlnAla	700
Db	2041	CATGAACAGGACCAATTTCAAGGACAGGTCGCCAGGGCATTAACCTTCACTGCTCCAAAGCA	2100
QY	701	GlnSerGlnProHisGlyValTyzProProSerSerValAlaGlnGlyGlnSerGlnGly	720
Db	2101	CAGTCTCAGCCCATGGGGTATATCTCAACCTTCMACTGTC-----	2139
QY	721	GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProLleGlnHisProGln	740
Db	2139	-----	2139
QY	741	GlnGlnGlnGlyLleGlnGlnThrAlaProProGlnGlnThrValGlnTySerLeuSer	760
Db	2140	---CAGCAGGGAAATACGACGACGACCCCTCTCTCAACGAGACGACGACTATTCCTTCA	2196
QY	761	GlnThrSerThrSerSerGlyAlaIleThrThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	2197	CAGACATCAACCTCCAGTAGAGGCCACTCTGCACAGCCAGTAGAGTAGCAGCTTCAAGCTTCA	2256
QY	781	GlnValLeuProGlnValSerAlaGlyLysGln-----	791
Db	2257	CAAGCTTGGCCTCAAGTATCAGCTGGAAAAACAGTTCCAGTTTCCAGCCAGTACCAACT	2316
QY	791	-----	791
Db	2317	ATCCAAGGCGAACCTCAATCCAGTTCGACACAAACCTCGGTTGTTCCAGTCACTCT	2376
QY	791	-----	791
Db	2377	GGTGTCAATTCTCTTCAGTGGAGACGCGCTCCCTATCTCCCTTGCTCCCTCAATACCT	2436
QY	791	-----	791
Db	2437	GTTCTCGAGATTCCTCATATCAATCTCCATGTGTCTACGCTCAGACAGAGTTTCTCATCC	2496
QY	791	-----	791
Db	2497	CTTCCCATCAGATGGAGCGAGCTGATTACTAGAGCTCTGCTCAGCTTGAGTTTCAATCTGCT	2556
QY	791	-----	791
Db	2557	ACAAAGCTGCAGATCCCGGGGTATCAACTGTGTCTTACTAGCTTCCAAACCTTCTG	2616
QY	791	-----	791
Db	2617	CAGCTGTGACTCAGCTGCCAAAGTCAGTTCAACCAAGCTCTCTTCAACCAAGCAGATTGAG	2676

QY	791	-----	791
Db	2677	TCCATTGGGAATACAGACTAACCTTGGACAAGCTGCTGAGTTTCACCTTCTGTGGAGAT	2738
QY	791	-----	791
Db	2737	GTTCTGTACAGSGCTCCCACTCGACTGCCACACAGTACCAGAGATTCAAATATT	2799
QY	791	-----	791
Db	2797	GCTCCCTCTTCCAAGTGAGCTTCTGTTCATTCATTCTACAGTCTCTAMCCCCCOCATG	2855
QY	791	-----	791
Db	2857	CCGACAGAAGTACTGGCTACACCTGGATACTTCCACAGTGTGCAGCTTAATGTGAA	2916
QY	791	-----	791
Db	2917	TCAATCTTTAGTTCCATTGGSTGTGAGAGACAGGTTCAAGTGTCCAGCCAGAA	2978
QY	792	-----SerThrGln	794
Db	2977	GGAGATTAGCACAGCCCCCACTCATCTCCCCAGACAAGCAGTTTGGAGAGTACTAG	3036
QY	795	GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln	814
Db	3037	GGAGTCTCTCAGGTGCTCTGTACAGCAGGACAGTTCAGTAGACACAGCCCAGCTACCGAG	3096
QY	815	ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer	834
Db	3097	CCGACCACTTTGGCTTCTCTGTACAGCTGCACATTGAGATGTTGCTTCAGGTATGAGT	3155
QY	835	AspGlyAsnGluAsnValProSerSerSerGlyArgHisGluGlyArgThrThryltyrArg	854
Db	3157	GATGGCAATGAGAACGTCCTCCATCTTCCAGTGGAGAGCATGAGAGAAACATCAAAAACG	3216
QY	855	HistylTyrglySerValArgSerArgSerArgHisGlysthrSerArgProIysLeu	874
Db	3217	CATTGCCAANAATCTGTAAAGAGTGCCTCTGCACATGAAAAAACTTCACGCCCAAAATTA	3276
QY	875	ArgIleLeuAsnValSerAsnLysGlyAspArgValValGluCysGlnLeuGlnThrHis	894
Db	3277	ACAAATTTGAAATGTTCAATTAAGAGAACCGAGTAGTAGAATGCAATVAGAGCTCAT	3338
QY	895	AsnArgLysmetValThrPheLysPheAspLeuAspGlyAspAsnProGluGlnIleAla	914
Db	3337	AATAGAAAATAGTACTTCAANTTCACTAGATGTGTACAACCCCGAGAGATACGA	3399
QY	915	ThrIleMetValAsnAsnAspPheIleuAlaIleGluArgGluSerPheValAspGln	934
Db	3397	ACAAATTAATGAGAACAAATGACTTTATTTTACCAATAGAGAGAGAGTCTTGTGTGATCAA	3456
QY	935	ValArgGluIleIleGluValAlaAspGluMetLeuSerGlyAspValSerValGluPro	954
Db	3457	GTCGAGAAATATTATTGAAAAGCTGATGAATGCTCATGTGAGAGTGTAGTGAACA	3516
QY	955	GluGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer	974
Db	3517	GAGGTTGATCAAGGATTTGGAAGTCTACAAGAAAGATAGACTATGTGTTTACAGTTCT	3576
QY	975	GlnLysLeuGluGluGluPheLysGlnProIleProAlaSerSerMetProGlnGlnIle	994
Db	3577	CAGAAATTTGGAAGGAGAGTTCMAAACCACTTCTGTCTTCATGTCCACACACAATA	3636
QY	995	GlyIleProThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgArgPheIleVal	1014
Db	3637	GGCAATCTTACAGATCTTTTAATCTCAAGATTGTATTCTCGGGAGAGCGGTTATATGTG	3696
QY	1015	SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluIleThrAsp	1034
Db	3697	AGTCTGTGCCAGAAAGCCGATTCAGAGAAATCAAAAGTTTTCCCCAGTGAATAACAGAT	3756
QY	1035	ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer	1054

```
Db 3757 ACAAGTGTGCTCTACAGCTCAGAGCCCTGGATGAACTGTCTCACTCTGCATCTATCC 3816
Qy 1055 LeuSerLeuGlnGlnAlaPheSerGlnLeuArgAlaGlnMetThrGlyProAsn 1074
Db 3817 CTGAGCTACAAAGAGGCTTTCTGAACTTAGACCTGCGCCAAATGACAGAGAGCCCAAY 3876
Qy 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValProProPheLeu 1094
Db 3877 ACAGACCTCCAAACTTAGTCATACAGACCCAAATTCAGATGTAACCTCTTCTTA 3936
Qy 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 3937 AGTAGCATGTGCTGAGGTCCCAACACAGACAGACACAGACAGCTCCCGAAACAGC 3996
Qy 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 3997 AGCCCTCTCAATGACATTCACATCAATTCAGTCTGAGGTACAGTCCCACTGAA 4056
Qy 1135 GluGlyIleAlaGlyValAlaAlaThrSerThrGlyValAlaThrSerGlyGlyLeuProIle 1154
Db 4057 GAGGGGATTCCTGAGTTCACACAGACAGAGTGTGTAATCTTCAGTGTCTCCCATTA 4116
Qy 1155 ProProValSerGlySerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 4117 CCACCTGTCTGATTCACACAGTACTTTCACAGCTGATTTCAAGTATCAACATATCTGCA 4176
Qy 1175 ValValSerIleSerThrThrSerProSerIleGlnValProThrSerThrSerGluIle 1194
Db 4177 GTTGCTCAATATCTACTACATCCCCCTCAGCTTCAAGTCCCAATCCCATCTGAGATC 4236
Qy 1195 ValValSerSerThrAlaLeuThrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 4237 GTTGCTTCTAGTACAGCATGTATCTTCAATACAGATTTACAGACACTTCCCTGCA 4296
Qy 1215 GlyGlySerThrAlaThrProGlyProIlyProIlyProAlaAlaValSerGlnGlnAlaAla 1234
Db 4297 GGGGGCAGTACTGCTACCCAGAGTCTTAAGCTCCAGCTGATGATTCAGACAGGACAGCA 4356
Qy 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 4357 GGCAGACACTACTGTGGAGCCACATTAACATAGTTTACACACACTTCATCCCAAGC 4416
Qy 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrIleuAlaGln 1274
Db 4417 ACAAGTTCACAGCTCTCAATTCAGCTTACAGCAATCTTCTACTCTTACTTACTGAA 4476
Qy 1275 ThrValValAlaSerAlaHisSerLeuAspIlyThrSerHisSerSerThrThrGlyLeu 1294
Db 4477 ACCGGTAGTGAAGGACACTCACTAGTAAAGACATCTCATAGCAGTAACTGATG 4536
Qy 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerIlyr 1314
Db 4537 GCTTCTCTCTCTGACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4596
Qy 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4597 ATTTCTACGCTGTGGGCTGTATCTTTGTGCTTTCATTCATCAGTATAGCTTCTACTCT 4656
Qy 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 4657 ATTCTCCCAAGAGCAGACGACCTTCTTCAACCTTTATATACCCCAAGTACTGATATC 4716
Qy 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 4717 CCACCTCTGTAACACCTGTTGCCAATGTGCTGTGTAACAGCAACCTAATTTATAGT 4776
Qy 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisIleGlyProGlnValAlaAspSer 1394
Db 4777 CAGCTCAACAGCTTTGCTTCCCAACAGCCCACTACTCATTTGTCTGAACTGATCTTCT 4836
Qy 1395 AspThrGlnProIlyAlaProGlyIleAspAspIleIleThrLeuGlnGlnIlyLeuArg 1414
Db 4837 GATACACACCCCAAGCTCTGTAATGATGATACATAAAGACTTCAAGAAAGAGCTGGC 4896
Qy 1415 SerLeuPheSerGlnHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 4897 TCTCTGTTCAGTGAACACAGCTCATCTGAGCTCAGACATGCTCTGTCTCATCTGAGACC 4956
Qy 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaAlaAlaProSer 1454
Db 4957 TCACTAGTATAGAGACATCTGTACACAGGACTCCCAACTTACTGCTGTTCACACAGC 5016
Qy 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 5017 AAATCTGACTTCTTACACACAGTACTGCTTACACACCAATTTACACTGAGAAACA 5076
Qy 1475 ValAlaLeuProValIleThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 5077 GTTCTTGTCCAGTATACACAGGTGTACACCTGAGCAAGTTCTTACCCAGTACACT 5136
Qy 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 5137 ACTACATCAGAGAGTGAACCTGGAACGTCTCTCCCAAGCACCCTTACTATAGGCTCG 5196
Qy 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnLeuProPro 1534
Db 5197 GTGCTGCAGAGGTGACTGAATTCACAGAGTACTTCAACAGCAGCAGCTGCACT 5256
Qy 1535 PheProGlyProSerLeuThrGlnSerGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 5257 TTTCAGAGCTTCTCTTACCCAGTCCACAGACCTCTAGAGGATCTGATGCTCAATG 5316
Qy 1555 ArgArgThrLeuSerProGly**IleThrValThrSerAlaValGlyProValSerMet 1574
Db 5317 AGAAGAACCTTACTATCCAGAGATKATACAGTACTTGTGGGTGTGCTGTGCAATG 5376
Qy 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnIlyGlyValSerGlnVal 1594
Db 5377 GCGGCTCCAAAGCAATCCACAGAGCAGAGCAAGCCTCAGAGGGGTTCCTCAATG 5436
Qy 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 5437 AAAGAGGCTCTGTCTTACAGACTATGTTACAGAGCTGTGTGTTTAAAGTGGAGCATTT 5496
Qy 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
Db 5497 CAGGTTCTGTTCAGACAGACAGGTGCCAAGAGGGTAAATAATGATCAGAAATGCA 5556
Qy 1635 LysSerValHisPheGlnSerSerThrSerGlnSerSerValLeuSerSerSerSerPro 1654
Db 5557 AAGTCTGTTCATTTGAAATCCAGACCTCAGAGTCTTCAAGCTATCAAGTAGTCCA 5616
Qy 1655 GluSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 5617 GAGGACTCTGTGTGAACCAAGCCGAAATGGCATACCACTCCGTGATCTCTTCAAT 5676
Qy 1675 ValProGlnSerAlaHisIlyThrThrAlaSerGlnAlaLysSerAspThrGlyGlnPro 1694
Db 5677 GTGCCAGAGAGTGCACCAAAACTACTGCTCAGAGGCAAGTCAAGACATGGGAGCCT 5736
Qy 1695 ThrLysValGlyArgPheGlnValThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 5737 ACCAAGGTGACGTTTTCAGGTGCAACTACAGAAACAAAGTGGTGTCTCTGTA 5796
Qy 1715 SerLysThrGlnAspLysIleThrAspThrLysLysGlnGlyProValAlaSerProPro 1734
Db 5797 TCAAAAACCTGAGACAGATCACTGACCAAGAAAGAAAGACAGTGGCATCTCTCT 5856
Qy 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlnLysProGln 1754
Db 5857 TTATGATTTGAAACAGAGCTTCTCTCTGTGTGATACCAAGAAAGAGAGAGCTGAA 5916
Qy 1755 LeuSerGlnProSerHisLeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSer 1774
Db 5917 CTGTCAAGGCTTCACTAATAATGGCGCTCTTCTGACCGGAGGCGGCTTTTAAAT 5976
```

QY	1775	ArgAspValAlaSpArgGlySerGlySerProHisSerProHisLeuLeuSerSerIysSer	1794
Db	5977	AGGGATGTGGAGTGAAGTGTCCGGTAGTCCACACTCGCCCACTACACTGACCTCAAGAGC	6036
QY	1795	LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSer	1814
Db	6037	CTTCTTACCAGAACTTAAGTCAAGACCTTAGTAACTTAATTAATTAATCTTCACTGAGT	6096
QY	1815	SerAspAsnGluSerAspIleGluAspGluAspLeuLysLeuGluLeuArgArgLeuArg	1834
Db	6097	AGCGCAATGAGTCAGATATCGAAGATGAAGACTTAAAGTTAGAGCTCGACACTGACA	6156
QY	1835	AspLysHisIleLeuLysGluIleGlnAspLeuGlnSerArgGlnLysHisGluIleGluSer	1854
Db	6157	GATTAACACTCTCAAGAGATTCAGGACTTCGAGAGTCCCGAAGACATGAATGAATCT	6216
QY	1855	LeuTyIThrLysLeuGlyLysValProProAlaValIleIleProProAlaIleProLeu	1874
Db	6217	TTGTATACCAACTGGGGCAGGTGCCCCCTGCTGTATTATTCCCCAGCTGCTCCCTT	6276
QY	1875	SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer	1894
Db	6277	TCAGGGAAACACGACGACCCACTAAAGCAAGCAGCAAAATCTAAGCGAAGAGTTCC	6336
QY	1895	LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaIleSerVal	1914
Db	6337	TTGGGGATTAAGAGCCCCCACTTTCAGGTAACTCGTGTGTGAGAGCGAGCTTCAGTC	6396
QY	1915	LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGluSerGlyGlnAsnGln	1934
Db	6397	TTGCACCCCCAGCAGACCTTCACCTCTCGCAACATCCAGAGTCGGGCAAGATACG	6456
QY	1935	LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuTyIserAlaPheThrSer	1954
Db	6457	CTGTTACGCCCCCTTAAGCCATCTCCCTCCAGTACAACTCTATTGAGCTTCACAGT	6516
QY	1955	AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn	1974
Db	6517	GATGTGTCATTTCAGTACCAAGCTTTCTCTCGAGGTCAAGAAACCGACGACCAAC	6576
QY	1975	ThrValGlyAlaThrValAsnSerGlnAlaIleGlnAlaGlnProProAlaMetThrSer	1994
Db	6577	ACTGTGGGGCAACAGTGAACAGCCCAAGCCCGCAAGCTCCTCGCATGACATCC	6636
QY	1995	SerIroLysGlyThrPheThrAspAspLeuHisLysLysLeuValAspAsnTTPAlaArgAsp	2014
Db	6637	AGCAGAAAGGCAATTCACAGATGACTTCGACAAAGTTGGTAGCAATTGGGGCCCGAGAT	6696
QY	2015	AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyIroGluIProGly	2034
Db	6697	GCATGAAATCTCTGAGCGAGAGAGAGAAAGCAAGGCACTGAATTTAGAGGGCCCTGGA	6756
QY	2035	MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly	2054
Db	6757	ATGGCAAGAAAGTTCTGTGACCTGGGGCAACTGTGCATCTCCAGTCACTCGAACC	6816
QY	2055	GlySerAlaProIleSerAlaAlaIleSerAlaThrSerLeuGlyHisPheThrLysSerMet	2074
Db	6817	GGCTCTCCCCCATCTCTGAGCATCAGCTACCTCTCTTAAGTCACTTCCACCAAGTCATG	6876
QY	2075	CysProProGlnGlnTyIroGlyPheProAlaThrProPheGlyAlaGlnTTPSerGlyThr	2094
Db	6877	TGCCCCCAGCAGCATGTGGCTTTCCAGCTACACCCCAATTTGGCGGTCAATGAGTGGAGC	6936
QY	2095	GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln	2114
Db	6937	GGTGGCCCAACACACGACCACTTGGCCAGTTCACCTGTGGGAATCGCTCCTCGTACG	6996
QY	2115	AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg	2134
Db	6997	AATTTCAACATCAGCAATTTGCAGAAATTCATCAGCAACCCCGAGCTCCTCAACTCGGG	7056

[illegible]

QY 261 GluArgGlnArgPheIysGluGluAlaGluMetLeuIysGlyLeuGlnHisProAsnIle 280
DB 601 GAGAGGACAGAGATTAAAGAAAGCTGAATGTTAAAGGTCTTCAGACATCCCAATATT 660
QY 281 ValArgPheIysAspSerIleIysGlnSerThrValIysGlyIysLysCysIleValLeuVal 300
DB 661 GTTAGATTATTATGATTCCTGGGAATCCACAGTAAAGAAAGAGTGCATTGTTGGTG 720
QY 301 ThrGluMetThrSerGlyThrIleuIysThrIleuIysSerPheIysValMetLys 320
DB 721 ACTGAACCTTAGACGTCTGGAACACTTAAACGATCTGAAAGGTTTAAAGTGTAG 780
QY 321 IleIysValIleuArgSerIleIysArgGlnIleuIysGlyLeuGlnPheLeuHisThr 340
DB 781 ATCAAAGTCTTAAGAGCTGGTGGCTGAGATCCTTAAAGCTTCACATTCCTTCATCT 840
QY 341 ArgThrProIleIleHisArgAspLeuIysCysAspAsnIlePheIleThrGlyPro 360
DB 841 CGAATCCACCTATCATTCACCGGATCTTAAATGTGACAAATCTTTATCACCGGCCCT 900
QY 361 ThrGlySerValIysIleGlyAspLeuGlyLeuAlaThrLeuIysArgAlaSerPheAla 380
DB 901 ACTGGCTCAGCAAGATTGAGACCTCGGTCTGGCAACCTGAAGCGGGCTTCTTTTGGC 960
QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetIleIysGluIleIysThr 400
DB 961 AAGAGTGTATGATGATCCCAAGATTCATGGCCCTGAGATGTATGAGAGAAATATATAT 1020
QY 401 GluSerValIleIysValIlePheGlyMetCysMetLeuGluMetAlaThrSerGluIle 420
DB 1021 GAATCCGTGTAAGTATTATGCTTTGGATGTGCATGCTTGAATGGTATCATCTGAAATAT 1080
QY 421 ProIysSerGluCysGlnAsnAlaAlaGlnIleIleIleIleIleIleIleIleIleIle 440
DB 1081 CCTTAAGTGTGAGTGTGCAAAATGCTGACAGATCTACCGTGGTGCACAGAGGGGTAG 1140
QY 441 ProIleSerPheAspIysValAlaIleProGluValIleIysGluIleIleGluGlyCysIle 460
DB 1141 CCAAGCAATTTGACAAAGTACATCTCTGATGAGAAAGAAATTTATGAAAGATGCATA 1200
QY 461 ArgGlnAsnIysAspGluArgIleIleIleIleIleIleIleIleIleIleIleIleIle 480
DB 1201 CGACAAACAAAGATGAAGAAATATTCATCAAAAGCCTTTTGAACCATGCTTCTTCCA 1260
QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspPheGlyGluIleIleAlaIle 500
DB 1261 GAGGAAACAGAGATCGGGTAGAATTTAGCAGAAAGATGATGAGAAAAATATACCATTA 1320
QY 501 LysLeuThrPheArgIleGluAspIleIleIleIleIleIleIleIleIleIleIleIleIle 520
DB 1321 AATATTATGCTACGATTTGAGATTTTAAAGAAATTTAAAGGAAATTAACAAATATATGA 1380
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet 540
DB 1381 GGTATTGAGTTTCTTTGATTGAGAGAGATGTCCCAAGAGATGTTGCAACAAGAAATG 1440
QY 541 ValGluSerGlyIleValCysGluGlyAspHisIleIleIleIleIleIleIleIleIle 560
DB 1441 GTAGAGTCTGGGTATGCTGTGTAAGTGTATCAACAAGCACTGCTAAAGCTTAAAGC 1500
QY 561 ArgValSerLeuIleIysArgIysArgGluGlnArgGlnIleuValArgGluGluGlnIle 580
DB 1501 AGAGATCTTAATTTAAGAGAAACAGAGCAGCGGAGTGTGTACGGAGAGAGCAAGAA 1560
QY 581 LysIleIysGlnGluGluSerSerLeuIysGlnGlnValGluGlnSerSerAlaSerGln 600
DB 1561 AAAAAAAGCAGAGAGAGAGAGCTCAACAAGCAGAGTGAACAATCCAGTCTTCCAG 1620
QY 601 ThrGlyIleIysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
DB 1621 ACAGAAATCAAGACAGCTCCCTTCTGTGTAGACACCGGATACCTACTGCTTCAACCTTCA 1680

QY 621 AlaSerValSerThrGlnValGluProGluGluIleAspGlnHisGlnGlnIleu 640
DB 1681 GCTTCAGTTCCTTACACAAGTGAAGTGAAGAACTTGAAGAGATCAACATCAACAACTA 1740
QY 641 GlnIleIysGlnIleProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnIle 660
DB 1741 CAGTACCAAGCAACCATATATCTGTGTATCTGATGAGAGGTTGACAGTGTACAGGA 1800
QY 661 SerSerValPheThrGluSerArgValSerSerGlnIleThrValSerIleIysGln 680
DB 1801 TCTCTGTCTTCAAGAAATCTCGATGAGCAAGCAACAGACAGTTCATATATGTTCCCA 1860
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGluHisIleProSerThrValGlnAla 700
DB 1861 CATGAACAGCAACATTTACAGGACAGTCCCAAGGAGATATACCTTCACTGTCCAAACA 1920
QY 701 GlnSerGlnProHisGlyValIleProProSerSerValAlaGlnGlnIysGlnIle 720
DB 1921 CAGTCTCAGCCCCAGGGGTATATCCACCTCAAGTGTGGCACAGGGGACAGGCCAGGT 1980
QY 721 GlnProSerSerSerSerIleuThrGlyValSerSerSerSerGlnProIleGlnHisPro 740
DB 1981 CAGCATCTCTCAAGTACCTTACAGGGGTTTCATCTTCCAAACCAATCAACATCTCAG 2040
QY 741 GlnGlnGlnIleIleGlnIleThrAlaProProGlnIleThrValGlnIleIysSer 760
DB 2041 CAGCAGCAGGGAATACAGCAGACAGCCCTCTCAACAGACAGTGCAGATATCACTTCA 2100
QY 761 GlnThrSerThrSerSerGluAlaIleThrAlaGlnProValSerGlnProGlnAlaPro 780
DB 2101 CAGACATCAACCTCAGTGAAGCCCACTGACAGCCAGAGAGTCAACCTCAAGCTTCA 2160
QY 781 GlnValLeuProGlnIleValSerAlaGlyIysGlnSerThrGlnIleValSerGlnValAla 800
DB 2161 CAGTCTTCTGCTCAAGTACCTGAGTGAAGAAACAGAGTACTCAGGAGTCTCAGTTCCT 2220
QY 801 ProIleGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
DB 2221 CCTGAGAGCCAGTTGAGTGAAGCAAGCCCAAGTACCCAGCCAGCACCTTGGCTTCC 2280
QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGlnAsnVal 840
DB 2281 TCTGTAGACAGTGCATCATTCAGATGTTCTCAGGTATGAGTGAAGTGAAGACATC 2340
QY 841 ProSerSerSerGlyIleArgHisGluGlyIleIleIleIleIleIleIleIleIleIleIle 860
DB 2341 CCATCTTCAGTGAAGGACATGAAGAAAGAACTACAAACCGCATTTACGAAATCTGTA 2400
QY 861 ArgSerArgSerArgHisGluIleIleIleIleIleIleIleIleIleIleIleIleIle 880
DB 2401 AGAGTGTGCTCTCGACATGAAGAAACCTTCAAGCCCAAAATTAAGAAATTTGATTTCA 2460
QY 881 AsnIysGlyAspArgValIleGluCysGlnIleuGluIleIleIleIleIleIleIleIle 900
DB 2461 AATAAAGAGACCGGATGTGAATGTCAATTTAGACTCTCAATTAAGAAATTTGTTTCA 2520
QY 901 PheIysPheAspLeuAspGlyAspAspProGluGluIleAlaIleIleIleIleIleIleIle 920
DB 2521 TTCAATTTTGAACCTGATGAGTGCACACCCGAGAGATGACATTAATGTGTAACAT 2580
QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
DB 2581 GACTTATTATTAAGCAATGAGAGAGATGCTTGTGATCAAGTCCGAGAAATTTTGA 2640
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGluIleAspGlnIleu 960
DB 2641 AAAGCTGATGAATGCTAGTGAAGATGTGGAATGTGAACAGAGGATGATGAGGATTTG 2700
QY 961 GluSerLeuGlnIleIysAspAspIleIleIleIleIleIleIleIleIleIleIleIleIle 980
DB 2701 GAGAGTCTACAGAGAAAGATACATATGCTTTTGAAGTCTCAAAATTTGAAAGAGAG 2760
QY 981 PheIysGlnProIleProAlaSerSerMetProGlnIleIleGlyIleProThrSerSer 1000

```
Db 2761 TTCAAAACCAATTCTCGCTTTCATGCCACAGCAATAGGCAATTCCTACAGTTCT 2820
Qy 1001 LeuThrGlnValValHisSerAlaGlyArgPheIleValSerProValProGlnSer 1020
Db 2821 TTAACGACAGTTGTTCTTCTCGGGAGAGCGGTTTATAGTACTCTGTCAGAGAAC 2880
Qy 1021 ArgLeuArgGlnSerIleValPheProSerGluIleThrAspThrValAlaIleSerThr 1040
Db 2881 CCAATTACGAGATCAAAAGTTTTCCTCCAGATGAATACAGATACAGTTCTGCTCTCA 2940
Qy 1041 AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnIleAla 1060
Db 2941 GCTCAGAGCCCTGGAGATGAGTCTGTCATCTGCATCATACCTTAAGTCTACAGAGGCC 3000
Qy 1061 PheSerGluLeuArgAlaGlnMetThrGlnGlyProAsnThrAlaProProAsnPhe 1080
Db 3001 TTTTGTGAACCTTAGAGCTGCCAAATGACAGAGAACCCAAATACAGCACTTCGAAATTT 3060
Qy 1081 SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGlyVal 1100
Db 3061 AGTCATACAGAGCAACCAATTCCTCAGTAGTACTCTTCTTAAGTAGCATTTGCTGGAGTC 3120
Qy 1101 ProThrThrAlaAlaIleThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
Db 3121 CCAACCAAGAGAGCCACAGACAGACAGCAGCTCCTGCAGCAAGAGCCCTCTATATGACATT 3180
Qy 1121 SerThrSerValIleGlnSerGluValThrValProThrGlnGluGlyIleAlaGlyVal 1140
Db 3181 TTCACATCAGTATATTCAGTCTGAGGTTACAGTCCCATGAGAGAGAGGATTCCTGAGATT 3240
Qy 1141 AlaThrSerThrGlyValValValThrSerGlyLeuProIleProProValSerGlnSer 1160
Db 3241 GCCACCAAGACAGAGGTGGTAACTTCAGGTGGTCTCCCATACCACTGTCGTCTGATACA 3300
Qy 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
Db 3301 CCAGTACTTCCAGGCTGAGTTCAGATCAACAAATACCTGCAGTGTCTCATATCTACTCT 3360
Qy 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla 1200
Db 3361 ACATCCCGCTCACTTCAGTCCCAATCCCATCTGATGATGATGTTCTTCTAGTACACA 3420
Qy 1201 LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
Db 3421 CTGATACCTCAGTAAAGTTTCAGCAACTTCAGGCTCTGCAGGGGAGATAGCTTACC 3480
Qy 1221 ProGlyProIleProProAlaValValSerGlnGlnAlaAlaGlySerThrValGly 1240
Db 3481 CCAGGTCCTAAGCTCCAGCTGTAGTATCTCAGAGGAGAGAGAGAGCACTACTGTGGGA 3540
Qy 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 3541 GCCACATTACATCAGTTTCTTACCAACCACTTCATTCCTCCAGAGCAAGCTTCACAGTGC 3600
Qy 1261 IleGlnLeuSerSerSerThrThrThrProThrLeuAlaGlnThrValValSerAla 1280
Db 3601 ATTCAAGCTTAGACAGAGTACTTCTACTCTTACTTAAAGTGAACCGTGATGATGAGCCA 3660
Qy 1281 HisSerLeuAspIleThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 3661 CACTCAGTAAATGAGACATCTCATAGCAGTACAACTGATGATGGCTTCTCTCCCTCTGCA 3720
Qy 1301 ProSerSerSerSerSerProGlyValAlaGlyValSerSerIleSerGlnProGlyVal 1320
Db 3721 CCACTTCTCTCTCTCTCTCTGAGAGAGAGATGATCTATTAATTTCTCAAGCTGGTGGG 3780
Qy 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
Db 3781 CTGCATCTTTGGTCAATTCATCAGTATGATGATCTTACTCTTATCTTCCCAAGAGCA 3840
Qy 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db 3841 GGACCTACTCTTACACCTTATATACCCCAAGTACCTAGTATCCACCCCTTGGTAGAGCT 3900
Qy 1361 ValAlaAsnValProAlaValGlnIleThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 3901 GTTGCCAAATGCTGCTGCTGACAGCAGACAACTTATCAATGACCTTCACAGCTTTG 3960
Qy 1381 LeuProAsnGlnProHisThrHisCysProGluValAlaAspSerAspThrGlnProIleAla 1400
Db 3961 CTGCCAACAGAGCCCACTACTATCTGCTTCAAGTAGTATTCGATACACACCCAAAGCT 4020
Qy 1401 ProGlyIleAspAspIleIleThrLeuGlnGlyIleLeuArgSerLeuPheSerGlnHis 1420
Db 4021 CCTGGAATTTGATGACATTAAGACTCTAGAGAAAGAGTGGGCTCTGTTGAGTAGAC 4080
Qy 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrSerLeuValIleGlnSer 1440
Db 4081 AGCTATCTGAGGCTCAGACATGCTCTGTCTCATCTGAGAGACTCATAGTCAATAGAGAC 4140
Qy 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerIleLeuLeuThrSerThr 1460
Db 4141 ACTGTCACACAGAGCATCCCAACTACTGCTGTCGACCAAGCAAACTCTGACTTCTTAC 4200
Qy 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4201 ACAAGTACTCTTCTTACCAACCAATTTACCATAGAGAAAGTTGCTTTGGCAGTTTACA 4260
Qy 1481 ProValValThrProProGlnValSerThrProValSerThrThrThrSerGlyValAla 1500
Db 4261 CCAGTGTCAACCTGGGAGAGTTCATCCCAAGTACACACTACTACATCAGAGATGAAA 4320
Qy 1501 ProGlyThrAlaProSerIleProProLeuThrIleAlaProValLeuProValGlyThr 1520
Db 4321 CTTGGAATGCTCCCTCCCAAGCCACCTTACTAAGCTCCGGTCTCCAGTGGTACT 4380
Qy 1521 GluLeuProAlaGlyThrLeuProSerGlnLeuProProPheProGlyProSerLeu 1540
Db 4381 GAACCTTCAGAGGACTCTTACCCAGCAGAGCTGCCACCTTTCCAGAGCTTCTCTA 4440
Qy 1541 ThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeuArgThrLeuSerPro 1560
Db 4441 ACCAGTCCCAAGCAACCTCTAGAGATCTTATGCTCAATGAGAGAAACCTTAGTCCA 4500
Qy 1561 Glu***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4501 GAGATKATCAAGAGACTTCTGGCGTTGTCTCTGTCTCAGAGGGCTCCCAAGCAATC 4560
Qy 1581 ThrGluAlaGlyThrGlnProGlnGlyValSerGlnValIleGlyGlyProValLeu 1600
Db 4561 ACAGAGCAGAGAACAGAGCTCAGAAAGGTGTTTCTCAAGTCAAAAGAGGCTGTCTTA 4620
Qy 1601 AlaThrSerSerGlyAlaGlyValPheIleMetGlyArgPheGlnValSerValAlaAla 1620
Db 4621 GCAACTGATGAGAGCTGGGTTTAAAGTGGAGCAATTCAGGTTTCTGTGACAGA 4680
Qy 1621 AspGlyAlaGlnIleGlyValAsnIleSerGlnAspAlaIleSerValIlePheGln 1640
Db 4681 GAGGTGCTCCAGAAAGAGGGTAAATTAAGTCAAAATGCAAAATGCTCTTCAATTTGAA 4740
Qy 1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValIle 1660
Db 4741 TCCAGCACTCAGAGCTCTCAGTCTCATCAAGTGTAGTCCAGAGAGTACTTGTTGAAA 4800
Qy 1661 ProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db 4801 CCAGAGCCGAATGAGATACCATCTCTGATATCTTCTCAAGTGGCCAGAGAGGCCAC 4860
Qy 1681 LysThrThrAlaSerGlnAlaIleSerAspThrGlnProThrIleValIleArgPhe 1700
Db 4861 AAAACTACTGCTCAGAGGCAAAATCAGACACTGGGAGCTTACCAAGTTTGACGTTTT 4920
Qy 1701 GlnValThrThrThrAlaAsnIleValGlyArgPheSerValSerIleThrGluAspIle 1720
Db 4921 CAGGTGACACTACAGCAAAACAAAGTGGTGGTCTCTCTGTATCAAAAACCTGAGAGCAAG 4980
```


QY 1121 ILeuThrAspThrIleuGluGlyProValAlaSerProProPheMetAspLeuGluGln 1740
 Db 4981 ATCACTGACACAAAGAAAGAGACAGTGGCATCTCTCTTTATGAAATTTGGAAACA 5040
 QY 1741 AlaValLeuProAlaValIleProLysLysGluLysProGluLeuSerGluProSerHis 1760
 Db 5041 GCTGTTCTTCTGCTGATACCAAGAAAGAAAGAAAGCTGAACTGTACAGACCTTCACAT 5100
 QY 1761 LeuSnglyProSerSerAspProGluAlaAphLeuSerArgAspValAspAspGly 1780
 Db 5101 CTAATGGGCGCTCTTGACCCGGAGCCGCTTTTAAAGAGAGATGTGATATGAT 5160
 QY 1781 SerGlySerProHisSerProHisGluLeuSerSerLysSerLeuProSerGluAsnLeu 1800
 Db 5161 TCCGGTATTCACACTGCGCCCATAGCTGAGCTCAAGAGCTTCCTAGCCAGAAATCTA 5220
 QY 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSerSerAspAsnGluSerAsp 1820
 Db 5221 AGTCAAAACCTTAGTAATTCATTTAACTCTTTCATGAGTAGAGCAATGACATGACAT 5280
 QY 1821 IleGluAspGluAspLeuLysLeuGluLeuArgArgLeuArgAspLysHisLeuLysGlu 1840
 Db 5281 ATCGAAGATGAAGACTTAAGTTAGAGCTGCGACGACTACGAGATAAACATCTCAAGAG 5340
 QY 1841 IleGluAspLeuGlnSerArgGlnLysHisGluIleGluSerLeuThrThrLysLeuGly 1860
 Db 5341 ATTCCGAGCTGACAGAGTGCACAGAGCATGAATGTAATCTTTGTAATCCAACTGGGCG 5400
 QY 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgArg 1880
 Db 5401 AAGGAGCCCTGCTGTATTTATTTCCCCAGCTGTCTCCCTTTCAGGAGAAAGAGAGCA 5460
 QY 1881 ProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerPro 1900
 Db 5461 CCCACTAAAGAGCAAGAGCAAGCAATCTAGTCGAAACAGATTCCTGGGAGATTAAGCCCC 5520
 QY 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
 Db 5521 CAGCTTTAGGTAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5580
 QY 1921 LeuHisProProGlyAsnIleProGluSerGlyGlnAsnGlnLeuLeuGlnProLeuLys 1940
 Db 5581 CTCCACCTCTCTGCGCAATCTCCCAAGTCCGGGAGAAATCAGCTGTTCAGCCCTTAAAG 5640
 QY 1941 ProSerProSerSerAspAsnLeuLysSerAlaPheThrSerAspGlyAlaIleSerVal 1960
 Db 5641 CCATCTCTCTCCAGTGAACAATCTATTCAGGCTTCACAGATGATGTGCCATTTCACTA 5700
 QY 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
 Db 5701 CCMAACCTTTCTGCTCAGGTAAAGAACACAGACAGCAACACATGTGGGGCAACAGTG 5760
 QY 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
 Db 5761 AACAGCCCAAGCCGCCAAGCTAGCTCTCTGCAAGTCCAGAGTCCAGAGAGAGGACATTC 5820
 QY 2001 ThrAspAspLeuHisLysLeuValAspAsnThrAlaArgAspAlaMetAsnLeuSerGly 2020
 Db 5821 ACAGATGACTTGCACAAAGTTGTAGACAATTTGGGCGCCGAGATGCCATGATCTTCCAGGC 5880
 QY 2021 ArgArgGlySerLysGlyHisMetAsnThrGlnGlyProGlyMetAlaArgLysPheSer 2040
 Db 5881 AGGAAGAGAGAGCAAGAGGACATGAATTTAGAGGCGCTCGAATGCAAGAGAGAGATTTCT 5940
 QY 2041 AlaProGlyGlnLeuLysLysSerMetThrSerAsnLeuGlyGlySerAlaProIleSer 2060
 Db 5941 GCACCTGGCACTGTGATCTCCATGACTCGAAGCTGGTGGCTCTCCCTCCCATCTCT 6000
 QY 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnThr 2080
 Db 6001 GGAGCATCAGTACTCTCTAGGTCACTTCAACCAAGTATGTGCCCCCAGACAGATAT 6060

QY 2081 GlyPheProAlaThrProPheGlyAlaGlnThrPheSerGlyThrGlyGlyProAlaProGln 2100
 Db 6061 GGCTTTCCAGCTTACCCCAATTTGGCGCTCAATGAGTGGAGCGGTGGCCCGACACACAG 6120
 QY 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
 Db 6121 CCACCTGGCCAGTTCCAACTGTGGGAACTGCTCTCTTCAAGAAATTTCAATCATCAGCAAT 6180
 QY 2121 LeuGlnLysSerLysSerAsnProProGlySerAsnLeuAlaGlnThr 2136
 Db 6181 TTGCAAGAAATCCATCAGCAACCCCGGCTCCCAACCTCGGAGACCAT 6228
 RESULT 10
 AR302607 6510 bp DNA linear PAT 12-JUN-2003
 LOCUS Definition Sequence 39 from patent US 6541252.
 ACCESSION AR302607
 VERSION AR302607.1 GI:31690905
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 REFERENCE
 1 (bases 1 to 6510)
 AUTHORS
 Walke, D.W., Hilpun, E., Donoho, G. and Turner, C.A. Jr.
 TITLE
 Human kinases and polynucleotides encoding the same
 JOURNAL
 Patent: US 6541252-A 39 01-Apr-2003;
 FEATURES
 Location/Qualifiers
 1..6510
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,81e-249 Length: 6510
 Score: 10451.50 Matches: 2076
 Percent Similarity: 95.71% Conservative: 0
 Best Local Similarity: 95.71% Mismatches: 0
 Query Match: 96.67% Indels: 93
 DB: Gaps: 1
 US-10-010-720-14 (1-2136) x AR302607 (1-6510)
 QY 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrThrGlnHisArgPhePhe 80
 Db 1 ATGACAAAGACACCGCGGGGGCGCGCCGCAACACTACACACATGACACACCGCTTCTTC 60
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 Db 61 CGCGGAGCGTCACTGTGCTGCTCCCAATGCTGCACTGAGAGCTTCCCGGCTTCTCTT 120
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
 Db 121 TCCCTGCCCCAGCCAGATCCCGGCGCTGTCGCGAGAGTGTCCACCGAGCCCGAC 180
 QY 121 ArgGlnGlnThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 181 CGGAAAGAGACCGTGCACCGCCACCTTCCAGGTAAGCCACAGACGCTTCCACCGCT 240
 QY 141 AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 241 GCCGCGCTCGGGGAAACAGGCGCTCGCGGGCGCTTCCCTGCACTGTCCCGACAGTGC 300
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProProPro 180
 Db 301 AGCAAAAGACCGCCAGTCTCCAGCTTGTGGGAGCAAGAGAGAGAGCGCGCGCG 360
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGlnProGlnGlnGluArgSerGlnGln 200
 Db 361 CGGAAGAATGGACAGGGGGGGGGGCGGCGCAAGAGCAAGAGAGAGAGAGAGAGAGAG 420
 QY 201 GlnAspAspIleGlnGlnLeuGlnThrLysAlaValAlaGlyMetSerAsnAspGlyArgPhe 220
 Db 421 CAGATGATATTCAGAGAGCTGAGACCAAGGCGGTGGAGATGTCTTAAGATGGCGCTTT 480

QY 221 LeuYsPheAspIleGluIleGlyArgGlySerPheIleThrValIleTyrIleGlyLeuasp 240
Db 481 CTCAGGTTGACATCGAATCGGACAGAGCTCTTTAAGACGGCTCTACAAAGGCTCGAC 540
QY 241 ThrGluThrValGluValAlaIleProCysGluLeuGlnAspArgIleLeuThrIleYsSer 260
Db 541 ACTGAAACCCACCCTGGAAGTCCCTGGTGTGAATCGAGGATCGAAATTTAACTAAAGCT 600
QY 261 GluArgIleArgPheIleGlyGluIleGluIleGluIleGluIleGluIleGluIleGluIle 280
Db 601 GAGAGCGAGAGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 281 ValArgPheTyrAspSerTyrCysIleThrValIleGlyIleGlyIleGlyIleGlyIle 300
Db 661 GTTAATATTTATGATTCCTCGGAATCCACAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 301 ThrGluLeuMetThrSerGlyIleThrLeuIleThrIleIleThrIleIleThrIleIle 320
Db 721 ACTGAACTTATGACGCTCGAAGACCTTAAAGGTAATCGAAAGGTTTAAAGTGAATGAAG 780
QY 321 IleIleValIleLeuArgSerTyrCysArgGlnIleLeuIleGlyIleGlyIleGlyIle 340
Db 781 ATCAAGTTCTAAGAGAGCTGGGCGCTGAGATCTTAAAGGCTCTTCAAGTTCTTCAATCT 840
QY 341 ArgThrProProIleIleHisArgAspLeuIleCysAspAsnIlePheIleThrGlyPro 360
Db 841 CGAATCTCCACCTATCATTCACCGCGATCTTAATGTGACCAACATCTTATCATCCGGCT 900
QY 361 ThrGlySerValIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGly 380
Db 901 ACTGGCTAGTCAAGATTTGAGAGACCTCGGCTCGGAGACCTGAAACCGGGCTCTTTTCC 960
QY 381 LysSerValIleGlyIleThrProGluPheMetAlaProGluMetTyrGluIleIleTyrAsp 400
Db 961 AAGAGTGTGATAGTACCCCAAGAGTTCATGGCCCTCGAGATGTATGAGAGAGAAATGAT 1020
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db 1021 GAATCCGTTGACGTTTATGCTTTGGAGTGTGATGCTTGAAGTGTGATGCTGATCATCTGA 1080
QY 421 ProTyrSerGluIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGly 440
Db 1081 CTTAATCTCGAGTGCAGAAATGCTGCACAGATCTAACCCTCGGACAGAGGGGTGAG 1140
QY 441 ProIleSerPheAspIleValAlaIleProGluValIleGlyIleGlyIleGlyIleGly 460
Db 1141 CCAGCCAGTTTTCACAAAGTGCATTCCTGAGTGAAGAGAAATTTATTTGAAGAGATGCA 1200
QY 461 ArgGlnAsnIleAspGluArgTyrSerIleIleIleIleIleIleIleIleIleIleIle 480
Db 1201 CGACGAAACCAAGATGAAAGATATTCATCAAGACCTTTGAACATGCTTCTTCCAA 1260
QY 481 GluGluThrArgValArgValGluLeuIleGluIleGluIleGluIleGluIleGluIle 500
Db 1261 GAGGAAACAGAGATGCGGTGAAATGAGACAGAAAGATGATGAGAGAAATTAATGCCATA 1320
QY 501 LysLeuThrPheLeuArgIleGluAspIleIleIleIleIleIleIleIleIleIleIle 520
Db 1321 AAATATATGCTACGATATTTGAAGATTTTGAAGATTTTGAAGAGAGAAATTTGAAGAT 1380
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGluIleMet 540
Db 1381 GCTATTAAGTTTCTTTGATTTAGAGAGAGATGTCCAGAAAGATGTTGACAGAAATG 1440
QY 541 ValGluSerGlyTyrValCysGluIleAspHisIleThrMetAlaIleAlaIleIleAsp 560
Db 1441 GTAGAGTCTGGATGCTGTGAAGGTGATCAAGAGCCATGGCTTAAAGCTATCAAGAGC 1500
QY 561 ArgValSerLeuIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 580
Db 1501 AGACTATCATTTAATTAAGAGAAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560

QY 561 LysIleIleGluIleGluIleGluIleGluIleGluIleGluIleGluIleGluIleGluIle 600
Db 1561 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 601 ThrGlyIleIleGluIleGluIleGluIleGluIleGluIleGluIleGluIleGluIle 620
Db 1621 ACAGGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 621 AlaSerValSerThrGluIleGluIleGluIleGluIleGluIleGluIleGluIleGlu 640
Db 1681 GCTTCAGTTTCTTCAAGATGAACTGAAAGACCTGAGGAGATCAACATCAACACTCA 1740
QY 641 GlnTyrGlnGlnProSerIleSerValIleSerAspGlyThrValAspSerGlyGlnGly 660
Db 1741 CAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 661 SerSerValPheThrGluSerArgValSerSerGlnIleThrValSerTyrGlySerGln 680
Db 1801 TCCCTGCTTCAAGAGATCTCGAGTGCAGAGCCAAAGAGAGAGAGAGAGAGAGAGAG 1860
QY 681 HisGluGlnAlaHisSerThrGlyIleThrValProGluIleIleIleProSerThrValGlnAla 700
Db 1861 CATGAACGACATTTCTACAGGACAGTCCAGGCAATATACCTTCTTACTGTCACAGCA 1920
QY 701 GluSerGlnProIleGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 1921 CAGTCTAGCCCATGGGATATTCACCCCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 1980
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 1981 CAGCATCTTCAGATGATTAACAGGGGTTTCATCTTCCCAACCATCAACATCTCTAG 2040
QY 741 GlnGlnGlnGlnIleGlnIleGlnIleThrAlaProGlnGlnIleThrValGlnIleSerLeuSer 760
Db 2041 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
QY 761 GlnThrSerThrSerSerGluAlaIleThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2101 CAGACATCAACCTCCAGTGAAGGCGCATCTGCAAGCGAGATGATGACCTTCAAGCTCCA 2160
QY 781 GlnValLeuProGlnIleValSerAlaGlyIleGlnIleGlnIleGlnIleGlnIleGln 791
Db 2161 CAACTCTTGCCTCAAGATACAGCTGGAAGAGAGGCTTCCACCTCGACTGACCCACAG 2220
QY 791 ----- 791
Db 2221 TACCAGAGATTAATATTTGCTCTCTTCCAGTGGCTTGTGTTGATTCATCTCT 2280
QY 791 ----- 791
Db 2281 ACAGTCTTAMCCCTCCCATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
QY 791 ----- 791
Db 2341 GTGGTGAAGCTTATGTGAATCAAAATCTTTAGTCTTATGGGTGTGATGAGAGAGAG 2400
QY 791 ----- 791
Db 2401 GTTCAAGTGTCCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
QY 792 -----SerThrGlnGlyValSerGlnValAlaProAlaGluProValAlaVal 807
Db 2461 GCAGTTTGGAGAGTACTCAGGAGGTCTTCAGGTTCTCTGAGAGAGAGAGAGAGAGAG 2520
QY 808 AlaGlnProGlnAlaIleThrGlnProIleThrLeuAlaSerSerValAspSerAlaHisSer 827
Db 2521 GCACAGCCCAAGGTACCCAGCCAGCACTTGGCTTCTCTGTAGACAGTGCACATTTCA 2580
QY 828 AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyIleArgHis 847
Db 2581 GATCTCTCTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
QY 848 GluGlyArgThrThrIleArgHisIleTyrArgIleSerValArgSerArgSerArgHisGlu 867

```
Db 2641 GAAGGAACAATCAAAACGGCATTAACGAAATTTGTAAGAGCGCTCTGCACATGAA 2700
Qy 868 LysThrSerArgProLysLeuArgLLeuLysnValSerAsnLysGlyAspArgValVal 887
Db 2701 AAAATTCACCCCAAAATTAAGATTGTAATGTTTCAAAATTAAGAGACCCGAGTATGTA 2760
Qy 888 GlyCysGlnLeuGluThrHisAspArgLysMetValThrPheLysPheAspLeuAspLys 907
Db 2761 GAATGTCAATTAGACATCATATATAGGAAATAGTTACATTCATAATTTGACCTAGATGCT 2820
Qy 908 AspAsnProGluGluLLeuAlaThrLLeuMetValAsnAsnAspPheLLeuAlaLLeuGlu 927
Db 2821 GACAAACCCGAGAGATAGCAACATTTATGTGAACATGACTTATTTCTAGCATATAG 2880
Qy 928 ArgGluSerPheValAspGlnValArgLLeuLLeuGlyLysPheLysGlnProLLeuProLa 947
Db 2881 AGAGAGTGTGTTGTGATCAAGTCCGAGAAATTAATTGAAAAAGCTGATGAAATGCTCAGT 2940
Qy 948 GluAspValSerValGluProGluGluLysAspGlnGlyLeuGluLysLeuGlnGlyLysAsp 967
Db 2941 GAGGATGTCAAGTGTGAACACAGAGGATGATCAAGGATTTGAGAGCTTCAAGAGAAAGAT 3000
Qy 968 AspTyrGlyPheSerGlySerGlnLysLeuGluGlyLysPheLysGlnProLLeuProLa 987
Db 3001 GACTATGCTTTTTCAGGTTCTCAGAAATTTGGAAGAGAGTTCAAAACCAATTCCTGCG 3060
Qy 988 SerSerMetProGlnGlnLLeuGlyLLeuProThrSerSerLeuThrGlnValValHisSer 1007
Db 3061 TCTTCCATGACCAACAGAAATAGGCAATTCCTACAGTTCTTTAACTCAAGTTGTTATCT 3120
Qy 1008 AlaGlyArgArgPheLLeuValSerProValProGluSerArgLysGlnLysSerVal 1027
Db 3121 GGGGAGAGCGGTTATATGTAGTCTGTCGAGAAAGCCATTTACGAGATCAAAAGATT 3180
Qy 1028 PheProSerGlnLLeuThrAspThrValAlaAlaSerThrAlaGlnSerProGlyLysAsn 1047
Db 3181 TTCCCAAGTAACATAACAGATACAGTTCGCTCTACAGCTCAAGCCCTGGAATGAAAC 3240
Qy 1048 LeuSerHisSerLysSerSerLeuSerLeuGlnGlnAlaPheSerGlnLeuArgArgAla 1067
Db 3241 TTGTCTCACTCTGCATCATCTCTTATGCTACAAACAGGCTTTTTCGAACCTTGAGAGCTGC 3300
Qy 1068 GlnMetThrGlnGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe 1087
Db 3301 CAATAGCAAGAGAGCCCAATACAGCACTCCAACTTTAGTCATACAGACCAACATTT 3360
Qy 1088 ProValValProProPheLeuSerSerLLeuAlaGlyValProThrThrAlaAlaLLeuThr 1107
Db 3361 CCAGTAGTAACCTCTTTCTTAAGTAGCATTCGTCGAGTCCCAACCAAGCAGCAGCAGCA 3420
Qy 1108 AlaProValProAlaThrSerSerProProAsnAspLLeuSerThrSerValLLeuGlnSer 1127
Db 3421 GCACAGAGCCCTCGCAACAAAGCAGCCCTCTATATGACATTTCCACATCAGTAATTCAGTCT 3480
Qy 1128 GluValLLeuValProThrGlnGluGlyLLeuAlaGlyValAlaThrSerThrGlyValVal 1147
Db 3481 GAGGTTACAGTCCCACTGAAAGAGGAGGATTCCTGAGTTGCCAAGCAGCAGCAGGTTGGTA 3540
Qy 1148 ThrSerGlyGlyLeuProLLeuProProValSerGluSerProValLLeuSerSerValVal 1167
Db 3541 ACTTCAGGAGTCTCCCACTACACACTGTCGATGATACCAAGTACTTCCAGCGTAGTT 3600
Qy 1168 SerSerLLeuThrLLeuProAlaValValSerLLeuThrThrSerProSerLeuGlnVal 1187
Db 3601 TCAGATATCAACAATCTGCAAGTGTCTCAATATCTACTACATCCCGCTCAGCTTCAAGTC 3660
Qy 1188 ProThrSerThrSerGluLLeuValValSerSerThrAlaLeuLysProSerValThrVal 1207
Db 3661 CCCAATTCACATCTGAGATGCTGTTTCTTATGACAGCACTTAATCCTTCAAGTAAACAGTT 3720
Qy 1208 SerAlaThrSerAlaSerAlaGlySerThrAlaThrProGlyProLysProProAla 1227
Db 3721 TCAGCAATTCAGCCTCTGACAGAGGAGGAGTACTGCTACCCAGAGTCTTAAGCTTCAGCT 3780
Qy 1228 ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer 1247
Db 3781 GTAGATCTCAGCGAGGAGCGAGGAGCACTACTGTGGAGGCACTTAATCATCACTTCT 3840
Qy 1248 ThrThrThrPhePheProSerThrAlaSerGlnLeuSerLLeuGlnLeuSerSerThrThr 1267
Db 3841 ACCACCACTTATTCACAGACAGCTTCACAGCTGTCCATTCAGCTTACAGCAGTACT 3900
Qy 1268 SerThrProThrLeuAlaGluThrValValSerAlaHisSerLeuAspLysThrSer 1287
Db 3901 TCTACTCTTACTTATGCTGAACACCGTGTAGTACGCACTCACTAGATAGACATCT 3960
Qy 1288 HisSerSerThrThrGlyLLeuAlaPheSerLeuSerAlaProSerSerSerSerPro 1307
Db 3961 CATAGCACTACAACTGATGAGCTTTCCTCTCTCTGACCAATCTTCTCTCTCTCT 4020
Qy 1308 GlyAlaGlyValSerSerTyrLLeuSerGlnProGlyLLeuHisProLeuValLLeuPro 1327
Db 4021 GGAGCAGAGTGTCTAGTATATTTCTCAGCTGTGGGCTGCACTCTTGTGCTATTCGA 4080
Qy 1328 SerValLLeuAlaSerThrProLLeuProGlnAlaAlaGlyProThrSerThrProLeu 1347
Db 4081 TCAGATAGACTTCTACTCTTATCTTCCCAAGCAGAGACCTTACTTACACCTTTA 4140
Qy 1348 LeuProGlnValProSerLLeuProProLeuValGlnProValAlaAsnValProAlaVal 1367
Db 4141 TTACCCCAAGTACCTAGATATCCACCTGTGACAGCTGTGTCCAAATGTGCTGTGTA 4200
Qy 1368 GlnGlnThrLeuLLeuHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr 1387
Db 4201 CAGCAGACACTAATTCATATAGTCAGCTCACACAGCTTTGCTTCCCAACAGCCCATCT 4260
Qy 1388 HisCysProGluValAspSerAspThrGlnProLysAlaProGlyLLeuAspAspLLeuLys 1407
Db 4261 CATGTCTCTGAAGATGATTGATACACAAACCTCTGGAATGATACATTAAG 4320
Qy 1408 ThrLeuGlnGluLysLeuArgSerLeuPheSerGlnHisSerSerSerGlyValAlaGlnHis 1427
Db 4321 ACTTAGAAGAAAGCTGGCTGCTGTTCAGTGAACAGCTCAATCTGAGACTGAGCAT 4380
Qy 1428 AlaSerValSerLeuGluThrSerLeuValLLeuSerThrValThrProGlyLLeuPro 1447
Db 4381 GCTCTGTCTCAGTGAAGACCTCACTAGTATAGAGACAGCTGACACCGAGCATCCA 4440
Qy 1448 ThrThrAlaValAlaProSerLysLeuLeuThrSerThrThrSerThrCysLeuProPro 1467
Db 4441 ACTACTGCTGTGACCAAGCAAACTCTGACTTCTACACAAAGTACTGTGTACACCA 4500
Qy 1468 ThrAsnLeuProLeuGlyLLeuValAlaLeuProValThrProValValThrProGlyGln 1487
Db 4501 ACCAATTTACCACTAGGAACAGTGTCTTGGCAGTTACACAGATGTACACCTGGGCA 4560
Qy 1488 ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys 1507
Db 4561 GTTTTACCCCAAGTACAGCACTATCATCAAGAGTGAACCTGGAACCTGCTCCCAAG 4620
Qy 1508 ProProLeuThrLysAlaProValLeuProValGlyThrGluLeuProAlaGlyThrLeu 1527
Db 4621 CCACCTTAACCTAAGCTCCGCTGTGCAAGTGTGCAAGTGTGCACTTCCAGCAGTACTCTA 4680
Qy 1528 ProSerGluGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnGlnProLeu 1547
Db 4681 CCAGCAGAGAGCTCAGACCTTTTCCAGGACCTTCTTAACCACTCCAGCAACCTCTTA 4740
Qy 1548 GluAspLeuAspAlaGlnLeuArgArgThrLeuSerProGluLLeuThrValThrSer 1567
Db 4741 GAGGATCTTGAATGCTCAATGGAAGAACACTTATGCCAGATATATACAGTACTCT 4800
Qy 1568 AlaValGlyProValSerMetAlaAlaProThrAlaLLeuGlnAlaGlyThrGlnPro 1587
Db 4801 GCGGTTGTCTCTGTCCATGTGCGGCTCCACAGCAATCACAGAAAGCAGGACAGGCT 4860
```

QY	1588	GlnYrsGlyValaSerGlnValaSerGlnGlyProValaLeuAlaThrSerSerGlyValaGly	1607
Db	4861	CbGAAAGGAGTGTCTTCbAGACTCAAGAAAGGCGCTGTCTTbAGCAACTbAGTTCbAGAGACTGTGT	4920
QY	1608	ValPheLysMeGlyArgPheGlnValaSerValaAlaAspGlyValaGlnLysGlnGly	1627
Db	4921	GTTTtTAAGATGGAGCGATtTCAGGTTCTGTtTGACAGACCGtTGCCAGAAAGAGGt	4988
QY	1628	LysAsnLysSerGlnAspAlaLysSerValHisPheGlnSerSerThrSerGlnSerSer	1647
Db	4991	AAAAATbAGTCAGAAAGATGCAAAAGCTGTTCATtTTGAATCCAGACCTCCAGAGTCCCA	5040
QY	1648	ValLeuSerSerSerSerProGlnSerThrLeuValLysProGlnProAsnGlyLeuThr	1667
Db	5041	GTGCATATCAAGATbAGTbAGTCCAGAGAGTbACTTGTGGAAACbAGGCCAAATGGCAATbAAC	5100
QY	1668	ILeProGlyLLeSerSerAspValProGlnSerAlaHisLysThrThrAlaSerGlnAla	1687
Db	5101	ATCCCTGGATbACTCTTCbAGATGTCCbAGAGTGGCCCAAAACTbACTbAGCTCCbAGAGCA	5166
QY	1688	LysSerAspThrGlnGlnProThrLysValGlyArgPheGlnValThrThrThAlaAsn	1707
Db	5161	AAAGTCbACACTGGGCACTbACCAAGGTTGGACCTTTbCAGGTbACbACbACbACbACbAAC	5220
QY	1708	LysValGlyArgPheSerValSerLysThrGlnAspLysLeuThrAspThrLysLysGln	1727
Db	5221	AAAGGGGAGTCTTCTCTGTATCAAAAACTbAGGACbAGbACbACTbGACbACAAAGbAAGA	5280
QY	1728	GlyProValAlaSerProProPheMetAspLeuGlnAlaValLeuProAlaValIle	1747
Db	5281	GGACCAAGGGATCTCTCTCTCTTTATbGATTTbGAAACAAGCTGTCTTCCbCGCTGTbATA	5340
QY	1748	ProLysLysGlnLysProGlnLeuSerGlnProSerHisLeuAsnGlyProSerSerAsp	1767
Db	5341	CCAAAGAAGAAGbAGbACTbGAACTbGACbAGCTTbCACbACTbTAAATGGCCGCTCTTbTbAC	5400
QY	1768	ProGlnAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHisSerPro	1787
Db	5401	CCGGAGGGCGCTTTTtTAAGTbAGGATbTGATbGTGGTbCCGGTbAGTCCbACbACTbCGCC	5466
QY	1788	HisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer	1807
Db	5461	CATbACbCTbAGACTCAAAAGbCCCTCTbTbCCbAGATbCTbAACTbCAAAAGbCTbTbAGTbATbTbCA	5520
QY	1808	PheAsnSerSerLysMetSerSerAspAsnGlnSerAspIleGlnAspGlnAspLeuLys	1827
Db	5521	TTTbAACTbCTCTTbCATbAGTbAGbGACbACAATbAGTbCATbTbCGAAAGATbGAAbGACTbTbAAG	5580
QY	1828	LeuGlnLeuArgArgLeuArgAspLysHisLeuLysGlnLileGlnAspLeuGlnSerArg	1847
Db	5581	TTbAGAGCTGGbACbAGbCTbAGbATbAAATbCTCAAAAGbATbTCbAGbACTbTbCAAGbCTbCAAGbCTbCGC	5640
QY	1848	GlnLysHisGlnLileGlnSerLeuLysThrLysLeuGlyLysValProProAlaValIle	1867
Db	5641	CAGAAGbACTbAAATbTbGATbCTTbTbGATbTbCCAAACTbGGGCAAGbTGbCCCTbCGCTbTbATT	5700
QY	1868	ILeProProAlaAlaProLeuSerGlyArgArgArgArgProThrLysSerLysGlySer	1887
Db	5701	ATtTCCbCCbAGCTbGTbCCCTbCTTbTbCAGGGAAGbAGbAGbACbCCbACTbTbAAAGbCAAAAGbGCAbGC	5760
QY	1888	LysSerSerArgSerSerSerSerLeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSer	1907
Db	5761	AAATbCTbAGTbCAAGbACbAGTbCTTbGGGGAATbTbAAAGbCCbCCbAGCTTbTbCAGGTbACbTbGTCT	5820
QY	1908	GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLeuHisProProGlyAsnLile	1927
Db	5821	GGTbCAGbAGTbGAGbCTTbCAGTbCTTbTbCACCbCCbCAGbAGbACbCTTbCCbACbCCCTbCTbGAGbCAACbTbC	5880
QY	1928	ProGlnSerGlyGlnAsnGlnLeuGlnProLeuLysProAspProSerSerSerAspAsn	1947
Db	5881	CCbAGbAGTbCCGGGCAbGATbCACTbGTbTbACbGCCCCCTbTbAAGbCCbATbCTbCCCTbCAGbTbAGbACbAC	5940

QY	1948	Leu1yrsSer1aPheThrsSerAsnG1yAla1leSerValProSerLeuSerAlaProGly	1967
Db	5941	CTCATTTCAGGCTTCACCAAGTAGGTCACATTTCAGTACCAAGCCTTTCTGCTCCAGCT	60000
QY	1968	GlnGlyThrsSerSerThrsAnthValG1yAlaThrValAsnSerGlnAlaGlnAla	1987
Db	6001	CAAGGAACACAGACACAAACACTGTTGGGCGCAAGTGAACAGCAAGCCGCCAACCT	6067
QY	1988	GlnProProAlaMetThrsSerSerArgLysGlyThrPheThrsAspAspLeuHisLysLeu	2007
Db	6061	CAGCTCCTCGCAACAGACGTCACAGAGAGAGGACATTCACAGATGACTTCACAAACTTG	61200
QY	2008	ValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGlyArgG1ySerLysG1yHis	2027
Db	6121	GTCAGCAATGGGCGCCAGATGCGCATGATCTTCAGGACGGAAGAGAGCAAGGGGAC	61800
QY	2028	MetAsnTrpGlnGlyProGlyLysMetAlaArgLysPheSerAlaProGlyGlnLeuCysIle	2047
Db	6181	ATGAATTATAGAGGCGCCCTGGAAATGGCAAGAAAGTCTCTGCACTGGGCAACTGTGATC	62400
QY	2048	SerMetThrsSerAsnLeuGlyGlySerAlaProIleSerAlaAlaSerAlaThrsSerLeu	2067
Db	6241	TCGATGACCTCGAAGCTGGGGTGGCTGCGCCCACTCTGACGATCAGCACTCACTCTTA	63000
QY	2068	GlyHisPheThrsSerMetCysProProGlnGlnTrpGlyPheProAlaThrProPhe	2087
Db	6301	GGTCACTTCACCAAGCTCTATGTGCCCCCACAGCAGATGTGCTTCCAGTACCCCAATT	63600
QY	2088	G1yAlaGlnTrpSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlnPheGlnPro	2107
Db	6361	GGCGCTCATGTGAATGGGAGCGGAGCGGCCACACACAGCACACTTGGCAGTTCCAACT	64200
QY	2108	ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsn	2127
Db	6421	GTGGGAACCTGGCTCCTTGACGAATTTCAACATCAGCAATTTGCAGAAATCATCATGCAAC	64800
QY	2128	ProProGlySerAsnLeuArgThrThr	2136
Db	6481	CCCCCAGGCTCCAACTGGCGGACCACT	6507
RESULT 11			
LOCUS	AB002342	6812 bp	mRNA linear PRI 10-MAY-2002
DEFINITION	Human mRNA for KIAA0344 gene, partial cds.		
ACCESSION	AB002342		
VERSION	AB002342.2	GI:20521006	
KEYWORDS	KIAA0344.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Nagase,T., Ishikawa,K., Nakajima,D., Ohira,M., Seki,N., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.		
TITLE	Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro		
JOURNAL	DNA Res. 4 (2), 141-150 (1997)		
MEDLINE	970349984		
PUBMED	9205841		
REFERENCE	2 (bases 1 to 6812)		
AUTHORS	Ohara,O., Nagase,T., Kikuno,R. and Nomura,N.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-MAR-1997) Osamu Ohara, Kazusa DNA Research Institute; 1532-3, Yana, Kisarazu, Chiba 252-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-552-3913)		
COMMENT	On May 9, 2002 this sequence version replaced gi:2280478. Sequence updated (22-Jul-1997).		
FEATURES	Location/Qualifiers		
SOURCE	1..6812		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		

```

/clone="hg01568"
/sex="male"
/clone_type="brain"
/clone_lib="pbluescriptII SK plus"
/notes="this sequence was replaced that of hg01486 cDNA as
a representative cDNA sequence for KIAA0344."
1..6812
/gene="KIAA0344"
<1..6203
/gene="KIAA0344"
/codon_start=3
/db_xref="gi:20521007"
/protein_id="BAA20802.2"
/translation="AATTTTTHRRFRSVCDSNAPALTEPLSLPOPSIAV
OSAPREPRERTATATTSOVAOPPAARBEQANAPATVTSSTKSPVOPS
LVGSEKPPPASGGGSAKPEERQODDIELTKAVGNSDRFLKFDIEIG
KGSFTVKKGLDTEITVAVWCELDRLTKSERQFEEMKLGLOHPNIVRYS
WSTVKGKCIIVLTMTSGTLTKTKVPMKIKVISMCRQILKLOELHTRP
IHRDKCDNIEITGTGSKVIGDGLATLTKASPAKVIIGPEMAEEMKEKES
VDVAFGCMLEMATSEYSPSCONAAQIVRVTSQVPSADKVAIDVEKIDGCI
RONKDERYSIKDILNHAFFOERTGVRLAEEDDEKAIKILMIRIENIDKIKGKVR
NEAIFSFDLERDVEDVAQEMVSEGYCEBDHKTMALIKRVSILKRBQOLVR
BEQKKKQESLKOQVQSSASQTGIQLPSASTGIPASTTSASVSTOVEPEEA
DOHQLOQVQPSISVLSQGVDSQSSVFESRVSQQTVSYGQHBQASHGTVP
HISTVQAQSQPHGVPPSVQAQSQSQSSSLTGSSSQPIQPOQOQIQCTAP
OOTVQSLSTSTSEATTAQVSOQAPQVLPVYSAKQSTQGVSOVAPAPVAVO
POATOPTLASSVDSHSDVASGMDGNBNPSSGRHGRHTTKHYKXSVRSRHE
KTSRKLAILVSNKGRDVEQLETHNRKMYTFKPDIDGNPEIATIMVNDITLA
TERSEFVQVEEIIKADEMSESVSVEPDQGLSEIQGKODVFGSQKLEKFKQ
PIPASMPQOIGIPITSLTOVVASRFRIVSPESRLREKVPSEIITDTVASTA
OSPGMNLSHSSLSLQOAFSELRAQTEGNTAPNPSHTGPFPVPVPSLISAG
VPTTAATAPAVATSSPNDISTSVIOSEVTPTEGAGVATSGVATSGGLPVPV
SESPLSSVSSITIPAVYASTTSPISOVPTSEIYVSATLPSVYATSTASAG
GSTAPGKPRPVAVSQAAAGSTTVATITSTSTSPFTSAQSLSTQSSSTSTTTLA
ETVVASHSLSLDTSHSSTGLAFSLASSSSSPGAVSYISQGPALHPLVPSVIA
STPLIPQAAPSTPLPQVPSIPPLVQVNAVPOQLIHSQGPALHPLPQPTHC
PEVSDTQPKAKAGIDIDIKLEKRLSEHSSGACASVLSLETSLVETVPGIP
TTAVAPSKLSTSTSTCLPPTNLPGTVALPVTVPVTPGQVSTPSTTSSGKPGTAP
SKPPLTKAPVLPGVTELPAGTLPSPROLPPPEGPSTLOSQPLEDIDRLTSPEMI
TVTSRVGVSMAAPPAITEAGTQPKVQVKEGVLATSSGAVGKMRFOVSVAADG
AOKGKNSSEDAKSHFESSTESSVSSPESTLVPEPGLITIPGISDVSPEAH
KTJASEAKSDTQPTKGRFOVTTTANVGRSVSKTDKIDTKKBPSPMPDL
EQAVLPAVLPKKEKELSEPSHLDNGSDPEAPAFSPHOSKSLSP
SONLSOSFNSFNSYSSDMSDEDLKLRLDKLKEILODSORKEHIES
LYTKIGKVPAYITIPPAISGRPRRPPSKGSKSPSSGKSPOLSGMISQOSA
SVLHROQLHPRGNIPESGQNLLOPLKPSRSDNLKSAFTSDCAISPSLSARQCT
SSTNVGATVNSQAAQAPPAWTSRKGTFTDHLKLVDMARDMNTSGRSGSHM
NYEGMARKSAPQOLCI SMTSNLGSAPISAAATSI GHTKSKMCPPOYGFPATP
FGAOWSGTGPAQPLQGFQPVGTASLQNFNISNLKSI SNPPGSLRMT"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 1,02e-248 Length: 6812
Score: 10435.00 Matches: 2065
Percent Similarity: 99.85% Conservative: 0
Best Local Similarity: 99.85% Mismatches: 1
Query March: 96.51% Indels: 2
DB: 9 Gaps: 2

```

US-10-010-720-14 (1-2136) x AB002342 (1-6812)

```

QY 69 AlaAlaThrThrThrThrGluHisArgPheArgSerValIleCysAspSer 88
Db 3 GCGGCGACCACTACCACTACCACTACCTTCCGCGGAGGCTCATCTGCCACTCC 62
QY 89 AsnAlaThrAlaLeuGluLeuProGlyLeuProLeuSerLeuProGlnProSerIlePro 108
Db 63 AATGCCACTGCACTGAGCTTCCCGGCTTCTTCTTCCCTGCCCAAGCCGACATCCC 122
QY 109 AlaAlaValProGlnSerAlaProProGluProHisArgGluGluThrValThrAlaThr 128
Db 123 GCGGCTGTCCCGCAGAGTCTCCACCGGAGCCCAACCGGAGAGAACCGGTGACCGCCACC 182

```

```

QY 129 AlaThrSerGlnValAlaGluGlnProProAlaAlaAlaProGlyGluGlnAlaVal 148
Db 183 GCGACTTCCAGAGAGAGCCAGAGGCTCCAGCGCTGCGCCCTGGGAAACAGGCGCTC 242
QY 149 AlaGlyProAlaProSerThrValProSerSerThrSerLysAspArgProValSerGln 168
Db 243 GCGGCGCTGCGCCCTCCAGTCTGCTCCAGCAGTACACCAAGACCGCCGAGTGTCCAG 302
QY 169 ProSerLeuValGlySerLysGluGluLeuProProAlaAlaArgSerGlySerGlyGly 188
Db 303 CCTACCTTGTGGGAGAGCAAGAGAGCGCGCGGAGAGAGGAGGAGCGCGCGCGCGC 362
QY 189 SerAlaLysGluProGlnGluGluArgSerGlnGlnAspAlaGluGluLeuGlu 208
Db 363 AGCCGCAAGAGCCACAGAGAGAGCGAGCCAGCGAGAGATTCGAAGAAGCTGGAG 422
QY 209 ThrLysAlaValGlyMetSerAsnAspGlyArgPheLeuLysPheAspIleGluIleGly 228
Db 423 ACCAAGCGCGGAGATGTCTAACGATGGCGCTTCTCAAGTTTGACATCGAAATCGGC 482
QY 229 ArgGlySerPheLysThrValTyrLysGlyLeuAspThrGluThrThrValGluValAla 248
Db 483 AGAGGCTCTTTAAGACGCTCTACCAAGGCTTGACACTGAAACCAACCGTGAAGTTCGC 542
QY 249 TrpCysGluLeuGlnAspArgLysLeuThrLysSerGlnArgGlnArgPheLysGluGlu 268
Db 543 TGGTGTGAAGTACGAGATCGAAATTAACAAGTCTGAGAGGCGAGATTAAAGAAAGA 602
QY 269 AlaGluMetLeuLysGlyLeuGlnHisProAsnIleValArgPheTyrAspSerTrpGlu 288
Db 603 GCTGAAAGTTTAAAGGCTTCAGCATCCCATATTTGTAATTTTATGATTTCCGGGAA 662
QY 289 SerThrValLysGlyLysLysCysIleValLeuValThrGluLeuMetThrSerGlyThr 308
Db 663 TCCACAGTAAAGGAAAGAGAGGAGCTTTTGGAGCTGAACTTATAGCGTGGAGAA 722
QY 309 LeuLysThrTyrLeuLysArgPheLysValMetLysIleLysValLeuArgSerTrpCys 328
Db 723 CTTAAGAGCTATCGAAGAAAGTTTAAGTATGAAGATCAAAAGTCTTAAGAAAGCTGGC 782
QY 329 ArgGlnIleLeuLysGlyLeuGlnPheLeuHisThrArgThrProProIleIleHisArg 348
Db 783 CGTCAGATCTCTTAAAGGCTTCAGTTCCTCATACCTCAACCTCCACATCATTCAGCC 842
QY 349 AsnLeuLysCysAspAsnIlePheIleThrGlyProThrArgLysSerValLysIleGlyAsp 368
Db 843 GATCTTAATGTGACAAATCTTTATCCAGCGCTCTAGCTGCTGATCAAGATTTGAGAC 902
QY 369 LeuGlyLeuAlaThrLeuLysArgAlaSerPheAlaLysSerValIleGlyThrProGlu 388
Db 903 CTCGGCTCGGCAACCTCGAAGGCGCTCTTTGCGAAGAGTGATAGTACCCCAAG 962
QY 389 PheMetAlaProGluMetTyrGluGluLysTyrAspGluSerValAspValTyrAlaPhe 408
Db 963 TTCAATGAGCGCCCTGAGATGATGAGAGAGAAATATGATGATCCGTGAAGTTATCTTT 1022
QY 409 GlyMetCysMetLeuGlnMetAlaThrSerGluTyrProTyrSerGluCysGlnAsnAla 428
Db 1023 GGGATGTGATGCTTGAGATGCTACATCTGATATCTCTTACTCGAGATGCCAAATGCT 1082
QY 429 AlaGlnIleTyrArgArgValThrSerGlyValLysProAlaSerPheAspLysValAla 448
Db 1083 GCGCAGATCTACCGTCCGCTGACAGTGGGGGGAAGCCAGGCTTTGACAAAGTACCA 1142
QY 449 IleProGluValLysGluIleIleGluGlyCysIleLeuGlnAsnLysAspGluArgTyr 468
Db 1143 ATTCTGAAGTGAAGAAATATTAAGAGATGACATACCAAAACAAAGATGAAGATAT 1202
QY 469 SerIleLysAspLeuLeuAsnHisAlaPhePheGlnGluGluThrGlyValArgValGlu 488
Db 1203 TCCATCAAGAGCTTTTGAACCATGCTCTTCCAAAGAGAAACGAGATACGGGATGAA 1262
QY 489 LeuAlaGluGluAspAspGlyGluLysIleAlaIleLysLeuThrLysArgIleGluAsp 508

```


Db 5637 TATTGAGCTTCAACAGTATGTCATTTCAGTACCAAGCTTTCTGCTCCAGTCA 5696
 Qy 1669 G1YThrSerSerThrAsnThrValG1AlaThrValAsnSerGlnAlaIaIaGln 1988
 Db 5697 GGAACACGACGACAAACACTGTTGGGCAACAGTGAACAGCCAGCCGCCAGCTCAG 5756
 Qy 1989 ProProAlaMetThrSerSerSerArgLysGlyThrPheThrAspAspLeuHisLysLeuVal 2008
 Db 5757 CCTCTGCGCATACGCTCCAGCAGGAAGGCGACATTCACAGATGACTTGACAAAGTTGGTA 5816
 Qy 2009 AspAsnThrAlaArgAspAlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMet 2028
 Db 5817 GACAAATTGGGCGCGAGATCCATGATCTCTCAGGCGAGAGAGAAAGGAGCAACAG 5876
 Qy 2029 AsnTYrGluGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuCysLysSer 2048
 Db 5877 AATTATGAGGGCGCTGGATGAGCAAGAGATTCTCTGCACTGGGCACTGTGACTCTCC 5936
 Qy 2049 MetThrSerAsnLeuGlyGlySerAlaProLysSerAlaIaIaSerAlaThrSerLeuGly 2068
 Db 5937 ATGACCTCGAACCTGGGTGGCTCTGCTGCCCATCTCTGAGCATCACTCTCTAGGT 5996
 Qy 2069 HisPheThrLysSerMetCysProProGlnGlnTYrGlyPheProAlaThrProPheGly 2088
 Db 5997 CACTTCACCAAGTCTATGTGCCCCCAGCAGAGATGGCTTCCAGCTAACCCCATTTGGC 6056
 Qy 2089 AlaGlnTrpSerGlyThrGlyGlyProAlaProGlnProLeuGlyGlyPheGlnProAl 2108
 Db 6057 GCTCAATGGAGTGGGAGCGGGTGGCCAGACACACAGCACTTGGCCAGTTCCACTGTG 6116
 Qy 2109 GlyThrAlaSerLeuGlnAsnPheAsnLysSerAsnLeuGlnLysSerLysSerAsnPro 2128
 Db 6117 GGAACCTGCTCTCTTCAGATTTTCAACATCAGCAATTGGCAAAATCCATCAGCAACCC 6176
 Qy 2129 ProGlySerAsnLeuArgThrThr 2136
 Db 6177 CAGGCTCCAACTCGGACCACT 6200

RESULT 12
 AR302604 6969 bp DNA linear PAT 12-JUN-2003
 LOCUS AR302604 Sequence 33 from patent US 6541252.
 ACCESSION AR302604
 VERSION AR302604.1 GI:31690902
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE
 1 (bases 1 to 6969)
 AUTHORS Walke, D.W., Hilbun, E., Donoho, G. and Turner, C.A. Jr.
 TITLE Human kinases and polynucleotides encoding the same
 JOURNAL Patent: US 6541252-A 33 01-Apr-2003;
 FEATURES Location/Qualifiers
 source 1..6969
 /organism="unknown"
 /mol_type="genomic DNA"

Alignment Scores:
 Pred. No.: 3.15e-247 Length: 6969
 Score: 10375.00 Matches: 2076
 Percent Similarity: 89.41% Conservative: 0
 Best Local Similarity: 89.41% Mismatches: 0
 Query Match: 95.96% Indels: 246
 DB: 6 Gaps: 1

US-10-010-720-14 (1-2136) x AR302604 (1-6969)

Qy 61 MetAspLysAspSerArgGlyAlaIaIaThrThrThrThrGluHisArgPhePhe 80
 Db 1 ATGGCAAGAGACAGCCGCGGGCGGCGCCGACCACTTACCACTGAGCACCGCTTTCTT 60

Qy 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 Db 61 CGCCGAGAGTCACTGGAGATCCATGCCATGCCACTGGAGACTTCCGGCTTCTCTT 120
 Qy 101 SerLeuProGlnProSerLysLeuProAlaAlaValProGlnSerAlaProProGluProHis 120
 Db 121 TCCCTGCGCCAGCCAGATCCCGCGGCTGTCCGAGAGTGTCTCACCGAGCCGAC 180
 Qy 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 161 CGGGAAGAGACCGTACCCGACCCGACTTCCAGTAAAGCCAGACGCTTCCAGCGCT 240
 Qy 141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 241 GCGGCGCCCTGGGGAACAGGCGCGTGGCGGCGCTGCCCCCTGCACGTCTCCGACAGTACC 300
 Qy 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProPro 180
 Db 301 AGCAAAAGACCGCCGAGTCTCCAGCTAGCCTTGTGGGAGCAAGAGAGAGCGCGCGCG 360
 Qy 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGlnGluAlaArgSerGlnGln 200
 Db 361 GCGAGAAAGTGGAGCGCGCGGCGAGCGCCAGAGGCAAGAGAGAAAGAGAGAGAGAG 420
 Qy 201 GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
 Db 421 CAGATGATATCGAAAGACTGGAAGCCAGGCGCGTGGGAATGTCTAAAGATGGCGGCTT 480
 Qy 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTYrLysGlyLeuAsp 240
 Db 481 CTCAAGTTGACATCGAAATCGGACAGAGCTCTTTAAGAGCGTCTCAAAAGTCTGGAC 540
 Qy 241 ThrGluThrThrValGluValAlaAlaTrpCysGluLeuGlnAspArgLysLeuThrLysSer 260
 Db 541 ACTGAACACCGCGGAGAGCTGCGCTGTGTGAATCGACGAGATCGAAATTAACAAAGCT 600
 Qy 261 GlnArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnLys 280
 Db 601 GAGAGGACAGAGATTTAAAGAAAGAGCTGAATGTAAAGATCTTCAACATCCCAATTT 660
 Qy 281 ValArgPheTYrAspSerTrpGluSerThrValLysGlyLysCysIleValLeuVal 300
 Db 661 GTTAGATTATATGATCTCTGGAAATCCACAGATTAAGAAAGAGATGATGTTTGGTG 720
 Qy 301 ThrGluLeuMetTrpSerGlyThrLeuLysThrTYrLeuLysArgPheLysValMetLys 320
 Db 721 ACTGAAGTATGAGAGCTCTGGAACCTTAAACGATCTGAAAGAGTCTTAAAGTATGTAAG 780
 Qy 321 IleLysValLeuArgSerTrpCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
 Db 781 ATCAAGTTCTAAGAGAGTGTGCGGTGCGTCAATCTTAAAGTCTTCAAGTTCTTCACTACT 840
 Qy 341 ArgThrProProLysLysLysArgAspLeuLysCysAspAsnLysPheIleThrGlyPro 360
 Db 841 CGAAGTCACTATCATTCACCGGAGTCTTAAAGTGAACAACATCTTATACCGCGCCCT 900
 Qy 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
 Db 901 ACTGGCTCAGTCAGATTTGAGAGACTCGGTGTGGCAACCTTGAAGGCGGCTTCTTTGCG 960
 Qy 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTYrGluGluLysTYrAsp 400
 Db 961 AAGAGTGTGATAGTACCCAGAGTTCATGCGCCCTGAGATGTAGAGAAATATGAT 1020
 Qy 401 GluSerValAspValTYrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTYr 420
 Db 1021 GAATCCGTTGACGTTTATGCTTTTGGAGTGTGATGCTTGAAGTGTCACTGCAATAT 1080
 Qy 421 ProTYrSerGluCysGlnAsnAlaIaGlnIleTYrArgArgValThrSerGlyValLys 440
 Db 1081 CTTTACTCGAGTCCAAAATGCTGCACAGATTTCACTGCGTGCAGACAGTGGGAGTGAAG 1140
 Qy 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460

Db 1141 CCAGCGACGTTTGACAAAGTACCAATTCCTGAGAGGAAAGAAATTTATGAGAGTGCATA 1200
 Qy 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuAsnHisAlaPheHesGln 480
 Db 1201 CGACAAACAAAGAGTAAAGATATTCATCAAGACCTTTTGAAACCATGCTCTTCCAA 1260
 Qy 481 GluGluThrGlyValArgValGluLeuAlaGluGluLysAspGlyGluLysIleAla 500
 Db 1261 GAGGAAACAGAGAGTACGGGTAGAAATTTAGCAGAAAGATGAGGAAATAATACCATTA 1320
 Qy 501 LysLeuTyrPheValArgIleGluAspIleLysLysLysLysGlyLysTyrLysAspAsnGlu 520
 Db 1321 AATATNTGCTACGATTCATTAAGAAATTTAAAGGAAATTAACAAAGATTAATGA 1380
 Qy 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 Db 1381 GCTATTGAGTTCCTTTGATTTAGAGAGATGTCCTCAAGAGATGTTGCACAAAGAAATG 1440
 Qy 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 Db 1441 GTAGAGTCTGGGTATGCTGTGAAGGTATATCAAGACCATGCTPAAGCTTCAAGAC 1500
 Qy 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGluGlnGlu 580
 Db 1501 AGAGATCATTTATTAAGAGAAACGAGAGCGCGGCACTTGCTACGGAGAGACAGAA 1560
 Qy 581 LysLysLysGluGluGluSerSerLeuLysGlnValGluGlnSerSerAlaSerGln 600
 Db 1561 AAAAAAAGCAGAGAGAGAGAGCTCTCAACAGAGGTAAAGAACTTCAGTCTCCAG 1620
 Qy 601 ThrGlyIleLysGluLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 Db 1621 ACGGAATCAACAGCTCTCTCTGTCAGACCGGATACCTACGCTTCAACACTTCA 1680
 Qy 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
 Db 1681 GCTTACGTTTCTACCAAGTAGAACCTGAAAGAACCTGAGGAGAGATCAACATCAACACTA 1740
 Qy 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 Db 1741 CAGTCCAGCAACCCAGTATATCTGTATATCTGATGAGACCGGTGACAGTGTGACAGGA 1800
 Qy 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 Db 1801 TCTCTGCTTCAAGAAATCTGAGAGACGACCAACAGACGTTTCATATGTTCCAA 1860
 Qy 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 Db 1861 CATGAACGAGGACATTTCTACAGCAACGTCCTCAAGGAGCATATCTTCTTCCACAGCA 1920
 Qy 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 Db 1921 CAGTCTCAGCCCCATGGGTATATCCACCTCAAGTGTGGCACAAGGGGAGGCAAGGCT 1980
 Qy 721 GlnProSerSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
 Db 1981 CAGCCATCTCAAGTAGCTTACAGGGGTTTCACTTCCCAACCATACAAATCTCTCAG 2040
 Qy 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
 Db 2041 CAGCAGCAGGGAATACAGCAGACAGCCCTCTCAACAGACAGAGTGAATTCACCTTCA 2100
 Qy 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 Db 2101 CAGACATCAACCTCCAGTGAAGGCACTACTGCAACAGCCAGTGAAGCTCAAGCTCA 2160
 Qy 781 GlnValIleuProGlnValSerAlaGlyLysGln----- 791
 Db 2161 CAAGTCTTGCCTCAAGTATCACTGGAAGAACGCTTCCAGTTCCAGCCAGTACCACT 2220
 Qy 791 ----- 791

Db 2221 ATCCAGGCGAACCCTCAGATCCAGTTGCGACACAAACCTCGTGTGTTCCAGTCACTCT 2280
 Qy 791 ----- 791
 Db 2281 GGTGCTCATTTCTTCCAGTGGGACAGCCGCTCCTACTCTCTTCTCTCCTCAGTACCT 2340
 Qy 791 ----- 791
 Db 2341 GTCCTCAGATTCCTATATCAATCTCTCATGTGTCTACGGCTCAGACAGTTTCTCATCC 2400
 Qy 791 ----- 791
 Db 2401 CTTCCCATCACAATGGCAGCTGATTAATCTCAGCTCTGTCTCAGCTGGCTTCACTGTCT 2460
 Qy 791 ----- 791
 Db 2461 ACAACGCTGCGATCCCGGGGTATCAACTGTGTCTTCTAGTCACTTCAACCTTCTGT 2520
 Qy 791 ----- 791
 Db 2521 CAGCTGTGACTCAGCTGCCAAGTCAGTTCAACCCAGCTCTTCAACAGCAGATTGAG 2580
 Qy 791 ----- 791
 Db 2581 TCCATGCGAATACAGAGTAACTTACCTTGAACAGCTGCTGAGTTCCACTTCTCTGAGAT 2640
 Qy 791 ----- 791
 Db 2641 GTTCTGTACAGGGCTTCCACCTTCACTGCTGCCACACAGTACCAGAGATTCAAAATAT 2700
 Qy 791 ----- 791
 Db 2701 GCTCCTCTTCCAGTGGCTTCTGTGTGATCCATTTCTACAGTCTTACCCCTCCATG 2760
 Qy 791 ----- 791
 Db 2761 CCGACAGAGTACGCTGCTACACCTGCTACTTCCACAGTGTGACAGCTTATGTGAA 2820
 Qy 791 ----- 791
 Db 2821 TCAATCTTTTATGTTCTTATGGGTGTGTGAGAGACAGTTCAAGTGTCCAGCCAGGA 2880
 Qy 792 ----- 792
 Db 2881 GGGAGTTTAGCAGCAAGCCCCCACTACATCTTCCAGCAAGCAGTTTGGAGGTACTCAG 2940
 Qy 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
 Db 2941 GGAATCTCTCAGGTGTCTCTGTCAGAGGACAGTTGACAGTGAAGCAAGCCCAAGCTTCCAG 3000
 Qy 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 Db 3001 CCGACACTTGGCTTCTCTGTGACAGTGCATTCAGATGATGTGCTTCAAGGATAGAT 3060
 Qy 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGluGlyValThrThrLysArg 854
 Db 3061 GATGCGATTAAGAAAGTCCCATCTTCCAGTGAAGGACATGAAGAAAGTCAAAACGG 3120
 Qy 855 HisTyrArgLysSerValArgSerArgSerArgHisGluLysThrSerArgProLysLeu 874
 Db 3121 CATTAACCAAAATCTGTAAAGAGTGTCTCTCAATGAAGAAATCTTCAAGCCCAAAATTA 3180
 Qy 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValAlaGluCysGlnLeuGluThrHis 894
 Db 3181 AGAATTTGAATGTTTCAATTAAGAGACGAGTGAATGATGATCAATTAAGACTCAT 3240
 Qy 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluGluIleAla 914
 Db 3241 AATAGGAATATGCTTACATTCATTTGACCTAGATGTGACAAACCCGAGAGATAGCA 3300
 Qy 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln 934
 Db 3301 ACAAATTATGTGAACAAATGACTTTTATCTAGCAATATGAGAGAGTGTGTTGTGATCAA 3360

```
QY 935 ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
Db 3361 GAGCGAGAAATATTGAAAAAGCTGATGAATGCTCAGTGAAGGATGTCTAGTGTGAGACCA 3420
QY 955 GluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
Db 3421 GAGGGGTGATCAGAGGATGAGAGTCTACAGAGAAAGATGATGCTTTCAGGTCTCT 3480
QY 975 GlnLysLeuGluGlyGluPheLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
Db 3481 CAGAAATTGGAGAGAGAGCTTCAACACCAATTCCTGCGTCTTCATGCGACAGACAAATA 3540
QY 995 GlyIleProThrSerSerLeuThrGlnValValHisSerAlaGlyArgArgPheIleVal 1014
Db 3541 GGCATTCTCACCAGTTCTTTAACTCAAGTTGTTCAATTCGCGGGAAGGGGTTTATAGTG 3600
QY 1015 SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluIleThrAsp 1034
Db 3601 AGTCTGTGGCCAGAAAGCCGATTACGAGAAATCAAAAGTTTCCCAAGTGAATTAACAGAT 3660
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db 3661 AAGTTGTCCTCTTACAGCTCAGAGCCCTGGAAATGAACTTGTCTCATCTTCATCATCC 3720
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgAlaGlnMetThrGluGlyProAsn 1074
Db 3721 CTTAGCTTCAACACAGGCCCTTTCTGAACTTAGACGCGCCCAATGACAGAAAGACCCAA 3780
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu 1094
Db 3781 AAGAGACCTCCCAACCTTATAGTATCAGAGCAACCAATTTCCAGTGTACTCTCTTCTTA 3840
QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 3841 AGTAGCATTTGCTGAGTCCCAACACAGAGACGACACAGACCAAGTCCCTGGAAACAAGC 3900
QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 3901 AGCCTCTCAATGACATTTCCACATCAAGTATTCAGTGTGAGGTTACAGTGTCCCATGTA 3960
QY 1135 GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLeuProIle 1154
Db 3961 GAGGGGATTTGCTGGAGTGGCCACGACAGAGTGTGTAACCTTCAGTGTGTCTCCCATTA 4020
QY 1155 ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 4021 CCACCTGTGTCTGAATTCACAGAGTACTTTCAGCGTATGTTCAAGTATCAACAATACCTGCA 4080
QY 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 4081 GTTGTCTCAATATCTACTACATCCCGGTCACTTCAAGTCCCAATCCACATCATGATGATC 4140
QY 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 4141 GTTGTCTTCTAGTACAGCACTGTATCTCTCAGTAAACAGTTTCAGCAACTTCAGCTCTGCA 4200
QY 1215 GlyGlySerThrAlaThrProGlyProLysProProAlaValValSerGlnGlnAlaAla 1234
Db 4201 GGGGGCAAGTACTGTAACCCCAAGTCTTAAGCCTCAAGCTGATGATTCACAGAGGACACA 4260
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 4261 GGCAGACACTACTGTGGAGCCACATTAACATCAAGTTTCTACACCACTTCATTCACAGC 4320
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrIleuAlaGlu 1274
Db 4321 ACAGCTTACAGCTGTCTCATTCAGCTTACAGCACTACTCTCATCTTACTTACTGTA 4380
QY 1275 ThrValValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu 1294
Db 4381 ACCGTGTAGTATTAGCGCACCTCACTAATTAAGACATCTCAAGACGTAACAAGTGGATTG 4440
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
Db 4441 GCTTCTCCTCTCTCGACACCATCTTCTCTTCTCTCTCGAGAGAGATGTCTACTATAT 4500
QY 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 4501 ATTCTCAGCCTGGTGGGCTGTGATCTTGTGATCTTCACATGATATACTTCTACTCTCT 4560
QY 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 4561 ATTCTTCCCAAGCAGAGAGACTTACTTCAACCTTTATTAACCCCAATACTCTGATATC 4620
QY 1355 ProProLeuValGlnProValAlaAlaAsnValProAlaValAlaGlnThrLeuIleHisSer 1374
Db 4621 CCACCTTGTGTAACACCTGTTCGCAATGTGCTGCTGTACAGACACACTAATTTACTACT 4680
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer 1394
Db 4681 CAGCTCAACAGACTTGTCTTCCCAACAGCCCACTACTCATTTGCTGAAGTATGATCT 4740
QY 1395 AspThrGlnProLysAlaProGlyLysAspAspIleLysThrLeuGluGluLysLeuArg 1414
Db 4741 GATACACAACCCAAAGCTCTGGAATTGATGACATAAAGACTTGAAGAAAGCTGGCG 4800
QY 1415 SerLeuPheSerGluHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 4801 TCTCGTTCAAGTAACACAGCTCATCTGAGCTCAGCATGCTCTGTCTTCACTGAGAC 4860
QY 1435 SerLeuValIleGluSerThrValThrProGlyLysProThrThrAlaValAlaProSer 1454
Db 4861 TCACATGATCATAGAGAGCACTGTCAACACAGGACATCCCACTAGCTGTGTGACCAAGC 4920
QY 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 4921 AAACCTCTGACTTCTTACCAAGTACTGTGTACCAACCAATTTACCACTAGAACCA 4980
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 4981 GTTGCTTGTCCAGTTACACCAAGTGTCTACACTGGGCAAGTTTCTACCCCACTCAGCAT 5040
QY 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 5041 ACTACATCAGAGGTAAGAACTGAAACTGCTCCCTCCCAAGCAACCTCTAATCAAGGCTCG 5100
QY 1515 ValLeuProValAlaGlyThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPro 1534
Db 5101 GTGCTGCCAGGTGACTGAACCTTCCAGCAGTACTTCAACCAAGACCACTGCCACT 5160
QY 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeu 1554
Db 5161 TTTCAGAGACTTCTCTTACCCCAAGTCCAGCAACTCTTACAGGATCTTGATGCTCAATTC 5220
QY 1555 ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValGlyProValSerMet 1574
Db 5221 AGAAGAACACTTAGTCCAGAGATATATCATCATGACTTCTGCGGTGTGCTGTGCTCATTC 5280
QY 1575 AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 5281 GCGGCTCAACAGCATATCAAGAACAGAACACAGCTCGAAGAGGGTGTTCCTCAATTC 5340
QY 1595 LysGluGlyProValLeuValAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 5341 AAAAGAGGCCCTGTGCTTACGAACTAGTTCAAGAGCTGTGTTTTAAATATGAGAGATT 5400
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGluGlyLysAsnLysSerGluAspAla 1634
Db 5401 CAGGTTTCTGTGAGCAGAGAGGTGCCAGAAAGGGGTAAATAATCAAGAGATGCA 5460
QY 1635 LysSerValHisPheGluSerSerThrSerGluSerSerValLeuSerSerSerPro 1654
Db 5461 AAGTGTGTCAATTTGAATTCAGCACTCAAGAGTCTCAAGTGTATCAAGTATGATGTCA 5520
QY 1655 GluSerThrLeuValLysProGluProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
```

```

Db 5521 GAGAGTACTGCTGTAACACGAGCCGATGCAACACCTCTGATCTCTCAAT 5580
Qy 1675 ValProGluSerAlaHisLeuThrThrAlaSerGluAlaLeuSerAspThrGlyPro 1694
Db 5581 GTGCAGAGAGTGGCCCAAACTACTGCTCAGAGGCAAAAGTCAGACACTGGGAGCT 5640
Qy 1695 ThrLeuValGlyArgPheGlnValThrThrThrAlaAsnLeuValGlyArgPheSerVal 1714
Db 5641 ACCAAGGTGGACGTTTCAGGTGACACTACAGCAAAAGAGTGGGTCTCTGTA 5700
Qy 1715 SerLeuThrGluAspLeuValLeuThrAspThrLeuGlyProValAlaSerProPro 1734
Db 5701 TCAAAACCTGAGGAGCAAGATCCTGACCAAAAGAAAGAGCAGCTGGACATCTCTCT 5760
Qy 1735 PheMetAspLeuGlnAlaValLeuProAlaValLeuProLeuGlyLeuProGlu 1754
Db 5761 TTTATGGATTGGAAACAAGCTGTTCTCTGCTGTGATACCAAAAGAAAGAGAGCTGAA 5820
Qy 1755 LeuSerGluProSerHisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSer 1774
Db 5821 CTGTACAGACCTTCACATCTAAATGGGCGCTTCTGTACCCGAGGCGGCTTTTAACT 5880
Qy 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLeuSer 1794
Db 5881 AGGATGTGGATGATGATGTTCCGCTAGTCCACACTGCCCCATCAGCTGCAAAAGAGC 5940
Qy 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSer 1814
Db 5941 CTCTCTAGCAGAAATCTTAAGTCAAAAGCCTTACTTAATTTCAATTTCTCTTACATGAGT 6000
Qy 1815 SerAspAsnGluSerAspLeuGluAspGluAspLeuValLeuGlyLeuArgPheLeuArg 1834
Db 6001 ACCGCAATGACTCGATATTCAGAGTGAAGCTTAAGCTTAAGCTTGCGAGCACTGACA 6060
Qy 1835 AspLeuHisLeuLeuGlnLeuGlnAspLeuGlnSerArgGlnLeuHisGlnLeuSer 1854
Db 6061 GATAAACAATCTCAAAAGAGATTGAGAGCTGCAAGAGTCCGCAAGACATGAATTTGATCT 6120
Qy 1855 LeuTyThrLeuLeuGlyLeuValProProAlaValLeuLeuProProAlaAlaProLeu 1874
Db 6121 TTGTTATCCAAACTGGGCAAGGTGCCCCCTGCTGTTATTTATTTCCCCAGCTGCTCCCTT 6180
Qy 1875 SerGlyArgArgArgArgProThrLeuSerLeuGlySerLeuSerSerArgSerSerSer 1894
Db 6181 TCAAGGAGAGACGACGACCCCTTAAGACAAAGGAGCAATCTAGTGCAGAGCACTTCC 6240
Qy 1895 LeuGlyAsnLeuSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
Db 6241 TTGGGGGAATTAAGAGCCCGAGCTTTCAGGTAACTGTGTGTCAGAGTGCAGCTTCAGTCC 6300
Qy 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnLeuLeuProGluSerGlyGlnAsnGln 1934
Db 6301 TTGCACCCCAACACCACTCCACCTCTGCAACATCCAGAGTCCGAGCAAGATCAG 6360
Qy 1935 LeuLeuGlnProLeuLeuProSerProSerSerAspAsnLeuTySerAlaPheThrSer 1954
Db 6361 CTGTTACAGCCCTTAAGCACTCTCCCTCAGTGAACACTTATTTACGCTTCACAGT 6420
Qy 1955 AspGlyAlaLeuSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
Db 6421 GATGTGGCATTTCAGTACCAAGCCTTTCTGCTCCAGGTCAAGAAACGAGCACAAC 6480
Qy 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSer 1994
Db 6481 ACTGTGGGGCAACAGTGAACGCAAGCCGCAAGCTCAGCTCTCCCTGACAGCTCC 6540
Qy 1995 SerArgLeuGlyThrPheThrAspAspLeuHisLeuLeuValAspAsnThrPheLeuAsp 2014
Db 6541 AGCAGAGAGGACCATTTCAAGATGACTTGCACAAAGTTGTTAGCAATTTGGGCGGAGAT 6600
Qy 2015 AlaMetAsnLeuSerGlyArgArgGlySerLeuGlyHisMetAsnTyGlnGlyProGly 2034

```

```

Db 6601 GCCATGATCTCTCAGGAGGAGGAGCAAAAGGACACATGAATATAGGAGGCCCTGGA 6660
Qy 2035 MetAlaArgLeuPheSerAlaProGlyGlnLeuCysLeuSerMetThrSerAsnLeuGly 2054
Db 6661 ATGGCAAGAAATCTCTGCACTGGGCAACTGGGCACTGCATCTCATACCTCGAACTGGAT 6720
Qy 2055 GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLeuSerMet 2074
Db 6721 GGCTCTGCCCACTCTGCAAGCATCAGCTACTCTTCTAGCTACTTCAACCAAGTCTATG 6780
Qy 2075 CysProProGlnGlnTyGlyPheProAlaThrProPheGlyAlaGlnTyPserGlyThr 2094
Db 6781 TGCCCCCAACAGAGTATGCTTTCAGCTACCTCCATTTGGGCTCAATGAGTGGAGC 6840
Qy 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
Db 6841 GGTGGCCCAAGACCAACAGCTGTCGCAATTCGAGTCCAGTGGAGTCTCTTGGAG 6900
Qy 2115 AsnPheAsnIleSerAsnLeuGlnLeuSerIleSerAsnProProGlySerAsnLeuArg 2134
Db 6901 AATTTCACATCAGCAATTTCAGAGAAATCCATCAGCAACCCCAAGCTCCCAACTGGGG 6960
Qy 2135 ThrThr 2136
Db 6961 ACCACT 6966

RESULT 13
AR302618 6147 bp DNA linear PAT 12-JUN-2003
LOCUS AR302618
DEFINITION Sequence 61 from patent US 6541252.
ACCESSION AR302618
VERSION AR302618.1 GI:31690916
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 6147)
AUTHORS
Walke, D.W., Hilbun, E., Donoho, G. and Turner, C.A. Jr.
TITLE
Human kinases and polynucleotides encoding the same
JOURNAL
Patent: US 6541252-A 61 01-Apr-2003;
FEATURES
location/Qualifiers
source
1..6147
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,596-246 Length: 6147
Score: 10344.00 Matches: 2048
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 0
Query Match: 95.67% Indels: 28
DB: Gaps: 1

US-10-010-720-14 (1-2136) x AR302618 (1-6147)
Qy 61 MetAspLeuAspSerArgGlyAlaAlaAlaThrThrThrThrGluHisArgPhePhe 80
Db 1 ATGGCAAGAAAGAGCGCTGGGGGGGCGCGCGACCACTTACCACTCAGAGCAAGCTTCTTC 60
Qy 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100
Db 61 CGCGGAGAGTATCTGGAGCTCCATGCACTGCACTGAGAGCTTCCGGGCTTCTCTT 120
Qy 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
Db 121 TCCCTGCCCAAGCCCAAGCATCCCGCGGTGTCGCAAGAGTCTCCACCGAGGCCCAAC 180
Qy 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
Db 121 TCCCTGCCCAAGCCCAAGCATCCCGCGGTGTCGCAAGAGTCTCCACCGAGGCCCAAC 180
Qy 141 AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160

```

```

Db      241 GCGGCCCTGGGAAAGGCGTCGGGGGCGCTCCCTCGACTGTCCCGAGTACC 300
Qy      161 SerIyAspArgProValSerGlnProSerLeuValGlySerIyGlnGluProPro 180
Db      301 AGCAAAAGACCGCCCACTGTCCAGCTTGTGGGAGCAAGAGAGCGCGCGCG 360
Qy      181 AlaArgSerGlySerIyGlyGlySerAlaIyGlyGluProGlnGluIyArgSerGln 200
Db      361 GGAGAGAGTGGCAGCGCGCGCGAGCGCCAAAGAGCCACAGAGAGAAAGAGCCAG 420
Qy      201 GlnAspAspIleGluGluIyLeuGluIyThrIyAlaValGlyMetSerAsnAspGlyArgPhe 220
Db      421 CAGGATGATATGAGAAAGCTGAGACCAAGCGCGGATGTCTAAACATGCGCGCTTT 480
Qy      221 LeuIyPheAspIleGluIleGlyArgGlySerPheIySerThrValIyIyIyGlyLeuAsp 240
Db      481 CTCMAAGTTGACATCGMAATCCGCGAGAGGCTCTTTAAGACGGTCTACAAAGGCTTGGAC 540
Qy      241 ThrGluThrThrValGluValAlaIyPheCysGluLeuGlnAspArgIyLeuThrIySer 260
Db      541 ACTGAAACACCGCTGGAAGTCCCTGGTGTGAACCTGAGAGATCGAAATTAACAAAGTCT 600
Qy      261 GluArgGlnArgPheIyGlyGluGluAlaGluMetLeuIyGlyLeuGlnHisProAsnIle 280
Db      601 GAGAGGCAAGATTTAAAGAGAAAGACTGAATGTTAAAGGCTTCAGCATCCCAATATT 660
Qy      281 ValArgPheIyAspSerThrIySerThrValIyIyGlyIyLeuCysIleValLeuVal 300
Db      661 GTTAATTTTATGATTCCTCGGAAATCCACAGTAAAGAAAGAAAGAGTCAATTTTGGTG 720
Qy      301 ThrGluLeuMetThrSerGlyThrLeuIyThrIyIyLeuIySArgPheIyValMetIyS 320
Db      721 ACTGAACCTTATGACCTGTGSAACACTTAAACGTAATCGAAAGGTTTAAAGTATGATGAG 780
Qy      321 IleIyValIleuArgSerThrCysArgGlnIleLeuIyGlyIyLeuGlnPheLeuHisThr 340
Db      781 ATCAAAAGTCTTAAAGAGCTGGTGGCGCTGAGATCTTAAAGGCTTCAGTTCTTCACT 840
Qy      341 ArgThrProProIleIleHisArgAspLeuIySAspAsnIlePheIleThrGlyPro 360
Db      841 CGAATCCACCTATCATTCATCCCGCAATCTTAAATGACACACATCTTATCAACCGGCGCT 900
Qy      361 ThrGlySerValIyIleGlyAspLeuGlyLeuAlaThrLeuIySArgAlaSerPheAla 380
Db      901 ACTGGCTCAGTCAGATTCGAGACCTCGGTCTGGCAACCTGAACCGGGCTTCTTTGCC 960
Qy      381 LysSerValIleGlyThrProGluPheMetAlaProGluMetIyGluGluIySArgAsp 400
Db      961 AAGAGTGTGATGATGATACCCCAAGATTCATGCGCCCTGAGATGTATGAGAGAAATATGAT 1020
Qy      401 GluSerValAspValIyThrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluIy 420
Db      1021 GAATCCGTTGACGTTTATGCTTTTGGAGATGTCATGCTTGAATGCTTCAATTAAT 1080
Qy      421 ProIySerGluCysGlnAsnAlaIleGlnIleIyThrArgValThrSerGlyValIyS 440
Db      1081 CTTTACTCGAGGTGCGCAAAATGCTGCACAGATCTACCGTCCGTACCAAGTGGGTGAAG 1140
Qy      441 ProAlaSerPheAspIyValAlaIleProGluValIyIyGluIleIleGluGlyCysIle 460
Db      1141 CCAGCGAGTTTGAACAAAGTACCAATCTCTGAAGGAAAGAAATTAATGAAGAGATGACATA 1200
Qy      461 ArgGlnAsnIySAspGluArgIySerIleIySAspLeuLeuAsnHisAlaPhePheGln 480
Db      1201 CGACAAACCAAGATGAAGATATTCATCAAAAGCCTTTTGAACCAAGCTTCTTCCAA 1260
Qy      481 GluGluThrGlyValArgValGluLeuAlaGluGluIyAspAspGlyGluIyIleAlaIle 500
Db      1261 GAGGAAACAGAGAGTACGGGTAGAAATTGACAGAAAGAGATGATGAGAAAAAATATGCCATA 1320
Qy      501 LysLeuThrPheLeuArgIleGluAspIleIySLeuIyGlyIySArgIySArgAspGlu 520

```

```

Db      1321 AAATTATGCTACGATTATGAAGATATTAAAGAAATTAAGGAAAAATACAAAGATATGAA 1380
Qy      521 AlaIleGluPheSerPheAspLeuGluIyArgAspValProGluAspValAlaGlnGluMet 540
Db      1381 GCTATTGAGTTTCTTTGATTTAGAGAGAGATGTCCCAAGAAAGATGTGCACAAAGAAAG 1440
Qy      541 ValGluSerGlyIyValCysGluGluIyAspHisIySThrMetAlaIySAlaIleIyAsp 560
Db      1441 GTAGAGTCTGGGTATGCTGTGAAGTGTATCAAGACCATGTCTAAAGCTTCAAAAGAC 1500
Qy      561 ArgValSerLeuIleIySArgIySArgGluIyArgGluIyLeuValArgGluGluGlu 580
Db      1501 AGAGTATCATTAATTAAGAGAAAGAGAGAGAGCGCGAGTTGTACGGAGAGCAAGAA 1560
Qy      581 LysIySArgGlnGluGluSerSerLeuIySArgGlnIyValGluGlnSerSerAlaSerGln 600
Db      1561 AAAAAAGCAGAAAGAGAGAGCTTCAACAGCAGAGTAAAGAAATCAAGTCTTCCAG 1620
Qy      601 ThrGlyIleIySArgGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrSer 620
Db      1621 ACAGGAATCAAGCAGCTCCCTTCTGTGACAGCGGCATACCTACTGCTTCAACACTTCA 1680
Qy      621 AlaSerValSerThrGlnValGluProGluIyProGluAlaAspGlnHisGlnGlnLeu 640
Db      1681 GCTTCAGTTTCTACACAAGTGAACCTGAAGAACTGAGGACAGATCAACATCAACAATA 1740
Qy      641 GlnIyArgGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db      1741 CAGTCCAGCAACCAAGATATCTGTGTATCTGATGAGGACGGTGTACAGTGTACAGGGA 1800
Qy      661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerThrGlySerGln 680
Db      1801 TCCTCTGTCTTACAGAAATCTCGAGTACAGCCAAAGACAGACAGATTCAATATGTTCCCA 1860
Qy      681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db      1861 CATGAACAGGACACTTCTACAGGACAGCTCCAGGGCATATACCTTCACTGTCCAAAGCA 1920
Qy      701 GluSerGlnProHisGlyValIyThrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db      1921 CAGTCTCAGCCCATGGGGATATTCACCTCAAGTGTG----- 1959
Qy      721 GlnProSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db      1959 ----- 1959
Qy      741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnIySerIySer 760
Db      1960 ---CAGCAGGGAATACAGCAGACAGCCCTCCTCAACAGACAGAGTATTCACCTTCA 2016
Qy      761 GlnHisSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db      2017 CAGACATTAACCTCCAGTGAAGGCCACTATGCAAGCCAGAGTACAGCTTCAAGCTTCA 2076
Qy      781 GlnValLeuProGlnValSerAlaGlyIySArgGlnSerThrGlnIyValSerGlnValAla 800
Db      2077 CAAGCTTGCTCCTCAAGTATACGCTGGAAGAAAGAGTATCAGGAGTCTCCAGGTTGCT 2136
Qy      801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
Db      2137 CCTGACAGCCAGATGTCAGTACAGCAGCCCAAGCTACCCAGCCGACCACTTGTGCTTCC 2196
Qy      821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
Db      2197 TCTGTAGACATGCATTCAGATGTGCTTCAAGTATGATGATGCAATGAGAAAGCTTC 2256
Qy      841 ProSerSerSerGlyArgHisGlnGlyIyArgThrThrIySArgHisIyArgIySerVal 860
Db      2257 CCATCTTCAAGTGAAGGACATGAAGAAAGAACTCAAAACCGCATTAACGAAATCTGTA 2316
Qy      861 ArgSerArgSerArgHisGlnGlyIyThrSerArgProIySLeuValGlyIleLeuAsnValSer 880
Db      2317 AGGAGTGGCTCTGACATGAAGAAACCTTCAAGCCCAAAATTAAGAAATTTGAATGTTTCA 2376

```

QY 881 AsnlysgIAspAspArgValValGluCysGlnLeuGluThrHisAsnArgLysMetValThr 900
Db 2377 AATAAAGAGAGACCGAGTAGTAAGTAATGCAATTAAGAGACCAATAAGGAAATGGATTACA 2436
QY 901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsn 920
Db 2437 TTCATATTTGACCTGATGATGTACAAACCCGAGAGATGCAACAATTATGTTGAACAAT 2496
QY 921 AspPheIleLeuAlaIleGluArgLysPheValAspGlnValArgGluIleIleGlu 940
Db 2497 GACTTATATCTGACATAGAGAGAGAGATGTTGTGATCAAGTCGAGAAATTTATTGAA 2556
QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGlyAspGlnGlyLeu 960
Db 2557 AAAGCTGATGAATGCTCAGTAGAGATGTCAGATGTGAACCAAGAGGTGATCAGGAGATTG 2616
QY 961 GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGluGlyGlu 980
Db 2617 GAGAGCTTACAGAAAGATGACTATGCTTTTCAAGTTTCAAGAAATTTGAAGAGAG 2676
QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
Db 2677 TTCAACAACCAATTCCTGCGCTCTTCATGCCACAGCAAAATGCGATTCTTCAAGTTCT 2736
QY 1001 LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGluSer 1020
Db 2737 TTAACTCAGATTGTTCTTCTGCGGAGAGCGGTTTATGTAGTCTCTGCGCAAGAAC 2796
QY 1021 ArgLeuArgLysSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
Db 2797 CGATTACGAGATCAAAAGTTTCCCGAGTGAATATACAGATACAGTTGCTGCTCTACA 2856
QY 1041 AlaGlnSerProGlyMetCysLeuSerHisSerAlaSerSerLeuSerLeuGlnAla 1060
Db 2857 GCTCAGAGCCCTGAGATGAACTTGCTCTCATCTGCATCATCCCTTAGCTTACAAACAGGCC 2916
QY 1061 PheSerGluLeuArgArgLaglnMetThrGluGlyProAsnThrAlaProAsnPro 1080
Db 2917 TTTTCTGAATCTAGAGTCCCAATGACAGAAAGGCCCAATACAGCACCTTCAACCTTT 2976
QY 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal 1100
Db 2977 ACTCATACAGAGCAACAATTTCCAGTAGTACCTCTTTCTTAGTAGATGCTGAGAGTC 3036
QY 1101 ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProAsnAspIle 1120
Db 3037 CCACCCACAGCAGCAGCCACAGCACAGTCCCTGCAACAGCAGCCCTCTATATGACAT 3096
QY 1121 SerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGlyVal 1140
Db 3097 TCCACATCTGATATTCAGTCTAGGTTACAGTGCCACTGAAGAGAGGATGCTGAGATT 3156
QY 1141 AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer 1160
Db 3157 GCCACACAGCAGAGGTGTGTAATCTCAGGTGTCTCCCATACCACTGTGTCTGATACA 3216
QY 1161 ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr 1180
Db 3217 CCAGACTTTCAGAGTAGTATTCAGATACCAATACCTGCAAGTGTCTCAATATCTACT 3276
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla 1200
Db 3277 ACATCCCGCTCACTTCAATCCCGACATCCACATCAGATGCTGTCTTCTCTGATACAGA 3336
QY 1201 LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlyLysSerThrAlaThr 1220
Db 3337 CTGTATCTCTCAGTAACGTTTACGCAACTTCAAGCTTGCAGAGGGGAGTACTGCTACC 3396
QY 1221 ProGlyProLysProProAlaValValSerGlnGlnAlaGlySerThrThrValGly 1240
Db 3397 CCAGGTCTTAACCTCCAGCTGATATCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCTG 3456

QY 1241 AlaThrLeuThrSerValSerThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
Db 3457 GCCACATTAACATCAGTTTCTTACACCACTTCAATCCCAAGCAGAGCTTCAAGCTGTCC 3516
QY 1261 IleGlnLeuSerSerSerThrSerThrProThrLeuAlaGluThrValValSerAla 1280
Db 3517 ATTCAGCTTAGAGAGATGACTTCTTCTCTCTTACTTATGCTGAACCGGTGTAGTACGCGA 3576
QY 1281 HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
Db 3577 CACTCACAATGAATGAACATCTCATAGCAGTACCACTGATGTGGCTTCTCCCTCTCTGA 3636
QY 1301 ProSerSerSerSerSerProGlyValGlyValSerSerTyrIleSerGlnProGlyGly 1320
Db 3637 CCATCTTCTCTTCTCTCTCTGAGACAGAGTCTTATTTATTTCTTCAAGCTGTGGG 3696
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
Db 3697 CTGCATCTTGTGATTCATCACTAGTAGCTTCTACTCTTATCTTCCCAACAGACA 3756
QY 1341 GlyProThrThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db 3757 GACCTACTTCTTACACCTTTATTTATCCCAAGTACTGATATCCACCTTGTGTACAGCT 3816
QY 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 3817 GTTGCCAAATGTGCGCTGTACAGCAGACATTAATCATATGACACCTCAACACCTTTG 3876
QY 1381 LeuProAsnGlnProHisThrHisCysProGluValAspSerAspThrGlnProLysAla 1400
Db 3877 CTTCACCAACAGCCCATATCTATGCTGCTGAGTAGATTTGATACCAACCACT 3936
QY 1401 ProGlyIleAspAspIleLysThrLeuGluLysLysLysLysLysLysLysLysLys 1420
Db 3937 CTTGAAATGATGACATGAAGCTCTGAAGAAAGAGCGGTCTCTGTCTGATGAAC 3996
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrSerLeuValIleGluSer 1440
Db 3997 AGCTATCTGAGAGCTCAGACATCCCTGTCTCAGCAGAGCCTCACTAGTACATAGAGAC 4056
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr 1460
Db 4057 ACTGCACACAGCAGGATCCCAACTACTGCTGTGACCAAGCAAACTCCTGACTTCAAC 4116
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4117 ACAGATCTTGTCTTACCAACCAATTTACCATGAGAAAGCTGTGTGCTGACGTAC 4176
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrSerGlyValLys 1500
Db 4177 CCAGGTGTACACCTGGGCAAGTTTCTACCCAGTCAAGCACTACATCAAGAGTGA 4236
QY 1501 ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr 1520
Db 4237 CCTGGAAGCTGCTCCCTCCAGGACACTTAACTAAGGCTCGGCTGCTGCACTGTGACT 4296
QY 1521 GluLeuProAlaGlyThrLeuProSerGluGlnLeuProProPheProGlyProSerLeu 1540
Db 4297 GAAGTTCAGAGAGGATCTTACCCAGCAGAGAGCTGCCACTTTTCCAGAACCTTCTCTA 4356
QY 1541 ThrGlnSerGlnProLeuGlnAspLeuAspAlaGlnLeuArgArgThrLeuSerPro 1560
Db 4357 ACCCAGTCCACAGCAACCTCTAGAGATCTGTATGCTCAATTGAGAAACACTTGTGCTCA 4416
QY 1561 Glu***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4417 GAGATATCACAGTACCTTCTGCGGTGTGCTGTGCTCAGTCCGCGCTCAACACCAATC 4476
QY 1581 ThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGluGlyProValLeu 1600
Db 4477 ACAGAAAGAGAGACAGCTCTAGAAAGGTGTTCTCAAGTCAAAAGAGCCCTGTCTCA 4536
QY 1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620

```

Db      4537  GAACTAGTTCAGAGAGCTGTTTAAAGATGGACGATTCAGGTTCTCTTCAGCA 4596
Qy      1621  AaSPGlyAlaGlnlySGLyLysAsnlySersGluAspAlaLysSeraIHisphGlu 1640
Db      4597  GACGGGGCCCAAGAGGGGTAAAAATTAAGTCAAGAGATGCAAAAGTCTTCAATTTGAA 4656
Qy      1641  SerSerThrsSerGluSerSerValLeuSerSerSerSerProGluSerThrsLeuValys 1660
Db      4657  TCCAGACACTCAGAGCTCAGTGCATCAGAGTAGTCCAGAGAGTACCTTGGTGA 4716
Qy      1661  ProGluProAsnGlyIleThrlleProGlyIleSerSerAspValProGluSeraIHis 1680
Db      4717  CCAGAGCCGAATGGATACCATCCCTGATCTCTTCAGATGTCCAGAGAGTCCAC 4776
Qy      1681  LysThrThrAlaSerGluAlaLysSerAspThrGlyGlnProThrllyValGlyProPhe 1700
Db      4777  AAAACTACTGCTTCAGAGGCAAGTACGACCTGGGAGCCCTAACAGGTTGGAGCTTT 4836
Qy      1701  GlnValThrThrAlaAsnLysValGlyArgPheSerValSerLysThrsGluAspLys 1720
Db      4837  CAGGTGACACTACAGCAAAACAAGTGGTCTTCTGTATCAAAAACAGAGACAG 4896
Qy      1721  IleThrAspThrLysLysGlyGlyProValAlaSerProProPheMetAspLeuGln 1740
Db      4897  ATCACTGACACAAAGAAAGAGAGACAGTGCATCTCTTATGATTTGAGATTGACAA 4956
Qy      1741  AlaValLeuProAlaValIleProLysLysGlyLysProGluLeuSerGluProSerHis 1760
Db      4957  GCTGTTCTTCTGCTGTGTATCAAGAAAGAAAGAGCCGAGACTGCAGAGCTTCACAT 5016
Qy      1761  LeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db      5017  CTAATAGGGCGCTTCTTGACCCGAGGCGCTTTTAAAGTAGGAGTGTGATGATG 5076
Qy      1781  SerGlySerProHisSerProHisGlnLeuSerSerLysLeuProSerGlnAsnLeu 1800
Db      5077  TCCGTTATGTCACACTCCGCCATCAGCTCAGACTCAAGAGCTTCACAGAGATCTA 5136
Qy      1801  SerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSerSerAspAsnGlnSerAsp 1820
Db      5137  AGTCAAAAGCTTAGTATTCATTATTCATTCCTTTCATGATGAGGACAAATGAGTCAAT 5196
Qy      1821  IleGluAspGluAspLeuLysLeuGluLeuArgArgLeuArgAspLysHisLeuLysGlu 1840
Db      5197  ATCGAAGATGAAGACTTAAGTTAGCTGCGACACGACTACAGATTAACATCTCAAGAG 5256
Qy      1841  IleGlnAspLeuGlnSerArgGlnLysHisGlyIleGluSerLeuTyThrLysLeuGly 1860
Db      5257  ATTCAGGACCTGCAGAGTCCGCGAAGCATGAAATTGATCTTGTATACCAAACTGGGC 5316
Qy      1861  LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgArg 1880
Db      5317  AAGGTGCCCCCTGCTGTTATTTATTCCTCCAGAGTGTCTCCCTTCAGGAGAAAGACGCA 5376
Qy      1881  ProThrllySerLysGlySerLysSerSerArgSerSerSerSerLeuGlyAsnLysSerPro 1900
Db      5377  CCCACTAAAGCAAGAGGACGCAAAATCTAGTGAACAGCTTCCTTGGGAAATAAAGCCCC 5436
Qy      1901  GlnLeuSerGlyAsnLeuSerGlyGlnSeraIAlaSerValLeuHisProGlnGlnThr 1920
Db      5437  CAGCTTTAGGTAACTCTGCTGTCAGAGTGCAGCTTCAAGTTCAGACCCCGACAGACC 5496
Qy      1921  LeuHisProProGlyAsnIleProGluSerGlyGlnAsnGlnLeuLeuGlnProLeuLys 1940
Db      5497  CTCACCTCTCTGGCAATCTCCAGAGTCCGCGGCAAGATCAAGCTTTTCAGGCCCTTTAG 5556
Qy      1941  ProSerProSerSerAspAsnLeuTySerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db      5557  CCATCTCCCTCCAGTGAACAACCTTATTCAGCTTCACACAGTGAAGTGCATTTCACTA 5616
Qy      1961  ProSerLeuSeraIaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980

```

```

Db      5617  CCAGACTTTTGTCTCCAGTCAAGAACACAGACACAAACACTGTTGGGCAACAGTG 5676
Qy      1981  AenSerGlnAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrsPhe 2000
Db      5677  AACAGCCAAAGCCGCGCAAGCTCAGCTCTGCTGCAAGACGTCCAGACGAAAGGACATTC 5736
Qy      2001  ThrAspLeuHisLysLysLeuValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGly 2020
Db      5737  ACAGATGACTTGCACAAATTTGGTAGCAATTTGGGCCGAGATGCTCATGATCTTCAGGC 5796
Qy      2021  ArgArgLysSerLysGlyHisMetAsnTyrgLysGlyProGlyMetAlaArgLysPheSer 2040
Db      5797  AGAAGAGAAACAAAGGCGACATGATATAGAGGCGCTCGAATGCAAGAAATTTCTCT 5856
Qy      2041  AlaProGlyGlnLeuCysLleSerMetThrSerAsnLeuGlyGlySerAlaProIleSer 2060
Db      5857  GCACCTGGCAACTGTGATCTCCATGACCTTCGAACTCGGGGTGGCTGCCCTCCATCTCT 5916
Qy      2061  AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnTy 2080
Db      5917  GCAGCATGACTACTCTCTAGTCACTTCACCAAGTCTATGTGCCCCCAGACAGAT 5976
Qy      2081  GlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThrGlyProAlaProGln 2100
Db      5977  GGCCTTCCAGCTACCCCATTTGGCGCTCAATGAGTGGAGCGGGTGCCAGCACACAG 6036
Qy      2101  ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
Db      6037  CCACCTGGCCAGTTCCAACTGTGGAACTCCTCTTCGAAATTTCAACATGACGCAAT 6096
Qy      2121  LeuGlnLysSerLleSerAsnProProGlySerAsnLeuArgThrThr 2136
Db      6097  TTGCAAGAAATCCATCAGCAACCCCGACGCTCCAACTCGGAGACCACT 6144

RESULT 14
AR302615 6426 bp DNA linear PAT 12-JUN-2003
LOCUS AR302615
DEFINITION Sequence 55 from patent US 6541252.
ACCESSION AR302615
VERSION AR302615.1 GI:31690913
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 6426)
AUTHORS
Walke, D.W., Hilbun, F., Donoho, G. and Turner, C.A. Jr.
TITLE
Human kinases and polynucleotides encoding the same
JOURNAL
Patent: US 6541252-A 55 01-Apr-2003;
FEATRES
Location/Qualifiers
source
1..6426
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 4,12e-245 Length: 6426
Score: 10287.50 Matches: 2048
Percent Similarity: 94.42% Conservative: 0
Best Local Similarity: 94.42% Mismatches: 0
Query Match: 95.15% Indels: 121
DB: Gaps: 2

US-10-010-720-14 (1-2136) x AR302615 (1-6426)
Qy      61  MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrGluHisArgPhePhe 80
Db      1  ATGACAAAGACACACCGTGGGCGCGCGCGACCACTCCACACACTGAGCAACGCTTCTTC 60
Qy      81  ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLysProLeu 100
Db      61  CGCGAGAGCGTCACTGTGAGCTCCAAATGCCACTGAGGCTTCCGGGCTTCTCTT 120
Qy      101  SerLeuProGlnProSerLleProAlaAlaValProGlnSeraIaProProGluProHis 120

```

Db	121	TCCCTCCCGCCAGCCAGCATCCCGCGGTGTCCCGCAGAGTCTCCACCGAGCCAC	180	Db	1201	CGACAAACAAAGATGAAATGATTCCATCAAGACCTTTGAACCATCTTCCCA	1260
Qy	121	ArgGluThrValThrAlaThrSerGlnValAlaGlnProProAla	140	Qy	481	GluGluThrGlyValArgValGluLeuAlaGluGluAspArgGlyGluLeuLeuLeu	500
Db	181	CGGGAAAGAGACCGTACCGCCACCCCTCCAGTAGCCACAGCTCCACCGCT	240	Db	1261	GAGGAAACAGAGTAACGGTAAATTACAGAAAGATGATGGGAAAAATACCAATA	1320
Qy	141	AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr	160	Qy	501	LysLeuTrpLeuArgGlyLeuAspGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	520
Db	241	GGCGCCCTGGGGAAACAGGCGGTGGGGCCCTGGCCCTCGACTGTCCCGCAGTACC	300	Db	1321	AAATTATGGCTACGATTGAAGATTTTAAGAAATTTAAAGGAAAAATACAAAGATTAATGA	1380
Qy	161	SerLysAspArgProValSerGlnProSerLeuValGlySerLeuGluGluProPro	180	Qy	521	AlaLeuGluPheSerPheAspLeuGluArgValProGluAspValAlaGlnGluMet	540
Db	301	AGCAAAAGACCGCCAGCTGTCCAGCTAGCTGTGGGGAGCAAAAGAGCGCGCGCG	360	Db	1381	GCTATTAGATTTCTTTGATTATTAAGAGAGATGTCACAGAAAGATGTTGACAAAGAAATG	1440
Qy	181	AlaArgSerGlySerGlyGlySerAlaGlyGluProGlnGluArgSerGln	200	Qy	541	ValGluSerGlyTyrValCysGluGlyAspHisIleThrMetAlaIleLeuAsp	560
Db	361	GCGAGAGTGGCAGCGCGCGCGCAGCGCCAAAGAGCAACAGAGCAACGAGCCAGCAG	420	Db	1441	GTAAGTCTGGGTATGCTGTGAAGTATCAAGACCAATGCTAAAGCTATCAAGAC	1500
Qy	201	GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAspAspGlyArgPhe	220	Qy	561	ArgValSerLeuLeuLeuArgLysArgGluGlnArgGlnLeuValArgGluGluGln	580
Db	421	CAGATGATATCAGAGAGCTGGAGACCAAGCGCGGAAATCTTACACATGGCGCTTT	480	Db	1501	AGAGTATCATTTAATTAAAGAGAAACAGAGCAGCGCGACTTGTTACGGAGAGCAGAGA	1560
Qy	221	LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp	240	Qy	581	LysLysLysGlnGluGluSerSerLeuLysGlnValGluGlnSerSerAlaSerGln	600
Db	481	CTCAAGTTTGACATGAAATCGCAGAGGCTCTTTAGACGGTCTACAAAGGTCTGAC	540	Db	1561	AAAAAAAGCAGAAAGAGCAGTCTCAACAGCAGTAGAAACAATCCAGTCTTCCAG	1620
Qy	241	ThrGluThrThrValGluValAlaTyrCysGluLeuGlnAspArgLysLeuThrLysSer	260	Qy	601	ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrSer	620
Db	541	ACTGAACCCACCGTGAAGTCCCTGGTGTGAATCGAGGATCGAAATTTAAACAAAGCT	600	Db	1621	ACAGAAATCAAGACACTCTCTTGTCTAGACCGGCAATACCTACTGCTTCAACCTTCA	1680
Qy	261	GlnArgGlnArgPheLysGluGlnAlaGluMetLeuLysGlyLeuGlnHisProAsnIle	280	Qy	621	AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu	640
Db	601	GAGAGCGCAGAGTTTAAAGAAAGCTGAATCTTAAAGTCTTTCAGCATCCCAATTT	660	Db	1681	GCTTCAGTTTCTACCAAGTAGAACTCAAGAACTGAGCGCAGATCAACATCAACAATA	1740
Qy	281	ValArgPheTyrAspSerTyrPheSerThrValLysGlyLysCysIleValLeuVal	300	Qy	641	GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly	660
Db	661	GTTAATTTTATGATTCCTCGGATCCACAGTAAAGGAAAGAGATGTTTGGTG	720	Db	1741	CAGTACCGCAACCCAGATATCTGTGTTATCTGATGGAGCGGTGACAGTGTACAGGA	1800
Qy	301	ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys	320	Qy	661	SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln	680
Db	721	ACTGAACCTATATACGCTCTGGAGACCTTAAACGTAATCGAAAGGTTTAACTGATGAAG	780	Db	1801	TCTCTGTCTTACAGAAATCTGAGTAGAGCAGCCAAAGACAGCTTCAATATGTTCCCA	1860
Qy	321	IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisIleThr	340	Qy	681	HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla	700
Db	781	ATCAAGTTCTTAAGAGCTGGTGGCTCGATCTTAAAGTCTTCAAGTTCTTCACTACT	840	Db	1861	CATGAACGAGCAATCTTACAGGACAGATCCAGAGGATATATCTTCACTGTCAGAGA	1920
Qy	341	ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro	360	Qy	701	GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly	720
Db	841	CGAACTCCACCTATCATTCACCGCATTTAAATGTGACAAATCATCTTATCACCGGCT	900	Db	1921	CAGTCTCAGCCCATGGGGTATATCCACCCTCAAGCTGTG	1959
Qy	361	ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla	380	Qy	721	GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740
Db	901	ACTGGCTCAGTCAAGATTTGGAGACCTCGGTGGCAACCTTAACCGGGCTCTTTTGGC	960	Db	1959	-----	1959
Qy	381	LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp	400	Qy	741	GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer	760
Db	961	AAGAGTGATAGTATGCCCAAGTTCATGGCCCTCGAGATGTATACAGAGAAATATGAT	1020	Db	1960	---CAGCAGGGAATACAGAGCAGCCCTCTCAACAGACAGTCAAGTATCACTTCA	2016
Qy	401	GlnSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr	420	Qy	761	GlnThrSerThrSerSerGlyAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	1021	GAATCCGTTGACGTTTATGCTTTTGGAGTGCATCTTGAAGTGGCTACATCTGAAT	1080	Db	2017	CAGACATCAACCTCCAGAGAGGCCACTACTGACAGCCAGTAGTACAGCTTCAACCTCCA	2076
Qy	421	ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys	440	Qy	781	GlnValLeuProGlnValSerAlaGlyLysGln	791
Db	1081	CCTTACTCGAGATGCCAAATCTGCAAGATCTACCGTGGTGAACAAGTGGGGAAG	1140	Db	2077	CAAGTCTTGGCTCAAGTATCACTGGAATAACAGGGCTTCCCACTCGACTGCACACAG	2136
Qy	441	ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle	460	Qy	791	-----	791
Db	1141	CCAGCCAGTTTGAACAAATAGCAATTCCTGAAGTCAAGAAATTTTAAAGATGCATA	1200	Db	2137	TACCCAGAGATTAATAATATGCTCCCTCTTCCAAAGTGGCTTGTGTCATCATCT	2196
Qy	461	ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln	480	Qy	791	-----	791
Db	1200	-----	-----	Db	2197	ACAGTCTTACCCCTCCATGCCAGAGAAATGATGCTGATACCTGGGTACTTCCACA	2256

Qy	791	-----	791	Qy	1108	AlaProValProAlaThrSerSerProProAsnAspIleSerThrSerValIleGlnSer	1127
Db	2257	GTGTGACAGCTTATGTGATCAATCTTTAGTTCTATGGGTGTGTAGAGAGACG	2316	Db	3337	GCACAGTCCCTGGACAAGACGCCCTCTATGACATTTCCACATCAGATTAATGAGTCT	3396
Qy	791	-----	791	Qy	1128	GluValThrValProThrGluGluGlyIleAlaGlyValAlaThrSerThrGlyValAla	1147
Db	2317	GTTCAAGTGTCCAGCCAGGAGGAGGTATTGACAAAGCCCACTACCTCCAGCA	2376	Db	3397	GAGTTTACAGTCCCACTGAAGAGGGGATTTGCTGAGTTGCCACAGCAGCAGGTGTGTA	3456
Qy	792	-----	792	Qy	1148	ThrSerGlyIleuProIleProProValSerGluSerProValLeuSerSerValVal	1167
Db	2377	GCAGTTTGGAGAGTACTCAGGAGAGTCTCAGTTGCTCTCCGACAGCAGTTGACAT	2436	Db	3457	ACTTGAAGTGTCTCCCAATCCACTGTGTGTGATACCAAGTCTTCCAGCGGTAGT	3516
Qy	808	AlaGlnProGlnAlaThrGlnProThrThrLeuAlaSerSerValAspSerAlaHisSer	827	Qy	1168	SerSerIleThrIleProAlaValValSerIleSerThrThrSerProSerLeuGlnVal	1187
Db	2437	GCACAGCCCCCAAGTCAACCCAGCCAGCACTTTGGCTTCTCTGTAGACAGTCCATTC	2496	Db	3517	TCAAGTACACATATCTCGCATTTCTCATATTTACTACATCCCGGTCACTTCAAGTC	3576
Qy	828	AspValAlaSerGlyMetSerAspGlyAsnGluAsnValProSerSerSerGlyArgHis	847	Qy	1188	ProThrSerThrSerGluIleValValSerSerThrAlaLeuThrProSerValThrVal	1207
Db	2497	GATGTGCTTCAGGTATGATGATGCAATGAGAAAGTCCCATCTTCCAGTGAAGGCAT	2556	Db	3577	CCCACTCCACATCTGAGATCGTTGTTCTAGTACAGCAGCTGTATCTTCAAGTAAAGT	3636
Qy	848	GluGlyArgThrThrLysArgHisArgArgLysSerValArgSerArgSerArgHis	867	Qy	1208	SerAlaThrSerAlaSerAlaGlyIleGlySerThrAlaThrProGlyProLysProProAla	1227
Db	2557	GAGGAAAGACACTCAAAAGCGCATTAACGAAATCTGTAAGAGTCTCTGACATGGA	2616	Db	3637	TCAGCACTTCAGCTCTGACAGGGGCGAGTCTGTACCCAGGTCTTAAGCTTCAGCT	3696
Qy	868	LysThrSerArgProLysLeuArgIleLeuAsnValSerAsnLysGlyAspArgValVal	887	Qy	1228	ValValSerGlnGlnAlaAlaGlySerThrThrValGlyAlaThrLeuThrSerValSer	1247
Db	2617	AAACCTTCAAGCCCAAAATTAGAAATTTGAATGTTTCAATTAAGAGACCGAGTATGA	2676	Db	3697	GTAAGTATTCAGCAGCAGCAGCAGCAGCTCTGTGGAGCCACTTAACATCAGTTTCT	3756
Qy	888	GluGlySerGlnLeuGluThrHisAsnArgLysMetValThrPheLysPheAspLeuAspGly	907	Qy	1248	ThrThrThrSerPheProSerThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThr	1267
Db	2677	GAATGTCAATTAAGACACTCATATAGAAATGTTTAACTCAATCAAAATTTGACCTATAGT	2736	Db	3757	ACCACACTTATTTCCCAAGACAGCTTCAAGCTGTCAATTCAGCTTACAGCAGTACT	3816
Qy	908	AspAsnProGluGluIleAlaThrIleMetValAsnAsnAspPheIleLeuAlaIleGlu	927	Qy	1268	SerThrProThrIleAlaGluThrValValSerAlaHisSerLeuAspLysThrSer	1287
Db	2737	GACAAACCCCGAGAGATAGCAACATTTATGTGAAACATGACTTATTTCTGCAATTAAG	2796	Db	3817	TCATCTCTACTTAACTGAAACCGGTGATTTAGCCGACACTCACTGATTAAGACATCT	3876
Qy	928	ArgGlnSerPheValAspGlnValArgGluIleIleGluLysAlaAspGluMetLeuSer	947	Qy	1288	HisSerSerThrThrGlyLeuAlaPheSerLeuSerAlaProSerSerSerSerPro	1307
Db	2797	AGAGAGTGTGTTGATGATCAAGTCCGAGAAATTTATGAAAAAGCTGAGTAATGTCTAGT	2856	Db	3877	CATGCACTTACATCGATTTGCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3936
Qy	948	GluAspValSerValGluProGluGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAsp	967	Qy	1308	GlyAlaGlyValSerSerTyrIleSerGlnProGlyGlyLeuHisProLeuValIlePro	1327
Db	2857	GAGGATGTGAGTGGAAACGAGAGGTGATCAGGGATTTGAGAGCTCTCAAGGAAAGAT	2916	Db	3937	GGAGCAGAGAGTGTAGTATATTTCTCAGCTGTGGGTGATCCTTTGGTCAATTC	3996
Qy	968	AspTyrGlyPheSerGlySerGlnLysLeuGluGlyGluPheLysGlnProIleProAla	987	Qy	1328	SerValIleAlaSerThrProIleLeuProGlnAlaAlaGlyProThrSerThrProLeu	1347
Db	2917	GACTATGCTTTTTCAGTTCTCAGAAATTTGAGAGAGAGTTCAACAAACCAATTCCTCG	2976	Db	3997	TCAGTATAGCTTCTACTCTCTATCTTCCCAAGCAGCAGGACCTACTTCTACACTTTA	4056
Qy	988	SerSerMetProGlnGlnIleGlyIleProThrSerSerLeuThrGlnValAlaHisSer	1007	Qy	1348	LeuProGlnValProSerIleProProLeuValGlnProValAlaAsnValProAlaVal	1367
Db	2977	TCTTCATGCCACAGCAATAGGCAATCTCTACCAAGTCTTTAACTCAAGTTGTTCAATCT	3036	Db	4057	TTACCCCAAGTACTAGTATCCACCTTGTGTACAGCTGTGGCAATGTGCTGCTGTA	4116
Qy	1008	AlaGlyArgArgPheIleValSerProValProGluSerArgLeuArgGluSerLysVal	1027	Qy	1368	GlnGlnThrLeuIleHisSerGlnProGlnProAlaLeuLeuProAsnGlnProHisThr	1387
Db	3037	GGGGAAAGCGGTTTATGAGTCTGTGCGAGAAAGCCCATTAACGAAATCAAAAGTTT	3096	Db	4117	CAGCAGACACTAATTAATGATGACAGCTCAACCACTTGTCTCCCAACAGCCCAATCT	4176
Qy	1028	PheProSerGluIleThrAspThrValAlaAlaSerThrAlaGlnSerProGlyMetAsn	1047	Qy	1388	HisCysProGluValAspSerAspThrGlnProLysAlaProGlyIleAspAspIleLys	1407
Db	3097	TTCCCGAGTGAATACAGATACAGTCTCTGCTCTACAGCTCAGAGCCCTCGAATGAC	3156	Db	4177	CATTGCTCTGAAGTAGATTTGATACCAACCCAAAGCTCTGTGAATTTGAAGACATAAG	4236
Qy	1048	LeuSerHisSerAlaSerSerLeuSerLeuGlnAlaPheSerGluLeuArgArgAla	1067	Qy	1408	ThrLeuGluGluLysLeuArgSerLeuPheSerGlnHisSerSerSerSerGlyAlaGlnHis	1427
Db	3157	TTGTCTCACTCTGCATCATCTCTTACCTCAACAGGCTTTTTCGAATTAAGAGTCC	3216	Db	4237	ACTCTAGAGAAAGCTGCGGTCTCTGTTCAGTGAACAGCTCATCTCGAGCTCAGAT	4296
Qy	1068	GlnMetThrGluGlyProAsnThrAlaProProAsnPheSerHisThrGlyProThrPhe	1087	Db	4297	GCCTCTGCTCACAGGAGACTCTCATGTATAGAGAGCACTGTACACACCGAGATCCCA	4356
Db	3217	CAATTGACAGAAAGACCAAAACAGCAGCTTCAAACTTTAGTCTATACGAGCAACAATTT	3276	Qy	1448	AlaSerValSerLeuGluThrSerSerLeuValIleGluSerThrValThrProGlyIlePro	1467
Qy	1088	ProValValProProPheLeuSerSerIleAlaGlyValProThrThrAlaAlaThr	1107	Db	4357	ACTACTCTGTGACCAAGCAAGCACTCTGATCTTCAACCAAGTACTGTGTTACACCA	4416
Db	3277	CCAGTAGTACTCTCTTCTTAAAGTACATGCTGTGAGTCCCAACCAACAGCAGCAGCACA	3336	Qy	1468	ThrAsnLeuProLeuGlyThrValAlaLeuProValThrProValValThrProGlyGln	1487

Db 4417 ACCAATTTACACCTAGAGACGCTTCTTCCCACTTACACAGCTGCTGAGCA 4476
 Qy 1488 ValSerThrProValSerThrThrThrSerGlyValLysProGlyThrAlaProSerLys 1507
 Db 4477 GTTCTACCCAGTCAGCAGTACTACATCAGAGAGAACTGGAACTGCTCCCTCCAG 4536
 Qy 1508 ProProLeuThrLysAlaProValLeuProValGlyThrGluLeuProAlaGlyThrLeu 1527
 Db 4537 CCACCTTACAAAGCTCCGCTGCTGCAAGTGGTACTGAACTTCCAGCAGTACTTA 4596
 Qy 1528 ProSerGluGlnLeuProProPheProGlyProSerLeuThrGlnSerGlnProLeu 1547
 Db 4597 CCGAGCAGACGCTGCCACTTTTCCAGAGACTTCTTCAACCCAGCTCCAGAGACTCTTA 4656
 Qy 1548 GluAspLeuAspAlaGlnLeuArgArgThrLeuSerProGlu**IleThrValThrSer 1567
 Db 4657 GAGGATCTTGATGCTCAATTGAGAAAGAACTTACTGACAGATATATACAGTACTTCT 4716
 Qy 1568 AlaValGlyProValSerMetAlaAlaProThrAlaIleThrGluAlaGlyThrGlnPro 1587
 Db 4717 GCGGTGCTCTGCTGCTCAGAGGCTCCCAACAGCAATCAGAGAGCAGAAACAGACT 4776
 Qy 1588 GlnLysGlyValSerGlnValLysGlnGlyProValLeuAlaThrSerSerGlyValAlaGly 1607
 Db 4777 CAGAAAGGTGTTTCTCAAGTCAAAAGAGCCCTGCTTCAAGCACTAGTTCAAGAGCTGCT 4836
 Qy 1608 ValPheLysMetGlyArgPheGlnValSerValAlaAlaAspGlyValGlnLysGlnGly 1627
 Db 4837 GTTTTAAAGATGGAGCGATTCAGATTCTGTTGACAGCAGAGCGGCCAGAAAGAGGCT 4896
 Qy 1628 LysAsnLysSerSerGluAspAlaLysSerValHisPheGluSerSerThrSerGluSerSer 1647
 Db 4897 AAAATTAAGTCAAGAGATGCAAGATCTGTTCAATTGAAATCCAGCACTTCAGAGCTCTCA 4956
 Qy 1648 ValLeuSerSerSerSerProGluSerThrLeuValLysProGluProAsnGlyIleThr 1667
 Db 4957 GTGCTATCAGAGTAAAGTCCAGAGAGTACTTGTGAAACCCAGAGCCGAGTGGCAATAC 5016
 Qy 1668 IleProGlyLysSerSerAspValProGluSerAlaHisLysThrThrAlaSerGluAla 1687
 Db 5017 ATCCCTGATCTCTTCAGATGTGCCAGAGAGTGGCCCAAAACATCACTGCTCAAGAGCA 5076
 Qy 1688 LysSerAspThrGlnGlnProThrLysValGlyArgPheGlnValThrThrAlaAsn 1707
 Db 5077 AAGTAGACACTGGGAGCTTCAAGAGTGGACGTTTCAAGCTACAACTACAGCAAC 5136
 Qy 1708 LysValGlyArgPheSerValSerLysThrGluAspLysIleThrAspThrLysLysGlu 1727
 Db 5137 AAAGTGGGTCGTTCTCTGATCAAAAACTGAGGCAAGATCATGACACAAAGAA 5196
 Qy 1728 GlyProValAlaSerProProPheMetAspLeuGlnAlaValLeuProAlaValIle 1747
 Db 5197 GGACCAAGGAGCATCTCCCTCTTATGATTGAAACAAGCTGTTCTCTGCTGTGATA 5256
 Qy 1748 ProLysLysGlnLysProGluLeuSerGluProSerHisLysAsnGlyProSerSerAsp 1767
 Db 5257 CCAAGAAAGAGAGAGCTGAACTGTCAAGCTTCAATCTMAATGGGCCGCTTCTGAC 5316
 Qy 1768 ProGluAlaAlaPheLeuSerArgAspValAspAspGlySerGlySerProHisSerPro 1787
 Db 5317 CCGGAGGCGGCTTTTAAAGTAGGAGATGTGATGATGTTCCGGTAGTCAACACCTGCC 5376
 Qy 1788 HisGlnLeuSerSerLysSerLeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSer 1807
 Db 5377 CATCAGCTGAGCTCAAAAGAGCTTCTTACCCAGAACTTAAAGCAAGCCTTGTATTTCA 5436
 Qy 1808 PheAsnSerSerLysMetSerSerAspAsnGluSerAspIleGluAspGluAspLeuLys 1827
 Db 5437 TTTAACTCTCTTAACTAGTACGACCAATGAGTCAAGATATGAAAGATGAAGACTTAAAG 5496
 Qy 1828 LeuGluLeuArgArgLeuArgAspLysHisLysLysGluIleGlnAspLeuGlnSerArg 1847

Db 5497 TTAGAGCTGGCAGCACTACGATATAAATCTCAAGAGANTTCAAGACTGAGAGTGGC 5556
 Qy 1848 GlnLysHisGluIleGluSerLeuThrLysLysLysValProProAlaValIle 1867
 Db 5557 CAGAGAGCATGAATGATCTTTGTATATCAAACTGGGCAAGGGGCCCTGCTGTATTT 5616
 Qy 1868 IleProProAlaAlaPheLeuSerGlyArgArgArgArgProThrLysSerLysGlySer 1887
 Db 5617 ATTCCTCCAGCTGCTCCCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5676
 Qy 1888 LysSerSerArgSerSerSerSerLysLysLysLysSerProGlnLeuSerGlyLysLeuSer 1907
 Db 5677 AAATCTAGTCAAGACAGTCTTCCAGGAGAAATAAAGCCCTTCAAGTAACTGCT 5736
 Qy 1908 GlyGlnSerAlaAlaSerValLeuHisProGlnGlnThrLysHisProProGlyLysIle 1927
 Db 5737 GGTCAAGATGAGCTTCACTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5796
 Qy 1928 ProGluSerGlyGlnAsnGlnLeuGlnProLeuLysProSerProSerSerAspAsn 1947
 Db 5797 CCAGAGTCCGGCAGAAATCAGCTGTACAGCCCTTAAAGCATCTCCCTCAGAGAGAGAG 5856
 Qy 1948 LeuThrSerAlaPheThrSerAspGlyAlaIleSerValProSerLeuSerAlaProGly 1967
 Db 5857 CTCTATTACGCTTCAACAGAGATGCTGCTCATTTCAATCAAGAGCTTCTGCTCCAGCT 5916
 Qy 1968 GlnGlyThrSerSerThrAsnThrValGlyValAlaThrValAsnSerGlnAlaGlnAla 1987
 Db 5917 CAAGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5976
 Qy 1988 GlnProProAlaMetThrSerSerArgLysGlyThrPheThrAspAspLeuHisLysLeu 2007
 Db 5977 CAGCTCTGCTGCAATGACGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6036
 Qy 2008 ValAspAsnThrPalaArgAspAlaMetAsnLeuSerGlyArgGlySerLysGlyHis 2027
 Db 6037 GTAGACATTTGGGCGCGAGATGCCATGATCTCTCAGCAGAGAGAGAGAGAGAGAGAGAG 6096
 Qy 2028 MetAsnLysGlnGlyProGlyMetAlaArgLysPheSerAlaProGlyGlnLeuLysIle 2047
 Db 6097 ATGAATTAAAGAGGCGCTGAGAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6156
 Qy 2048 SerMetThrSerAsnLeuGlyGlySerAlaProIleSerAlaAlaSerAlaThrSerLeu 2067
 Db 6157 TCCATGACCTGAGACCTGAGGAGGCTGCTGCTCCCACTCTGAGAGATCAGTACTCTCTA 6216
 Qy 2068 GlyHisPheThrLysSerMetCysProProGlnGlnIleThrGlyPheProAlaThrProPhe 2087
 Db 6217 GGTCACTTCAACCAAGTCTATGTGCCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6276
 Qy 2088 GlyAlaGlnThrPheSerGlyThrGlyValProAlaProGlnProLeuGlyGlnPheGlnPro 2107
 Db 6277 GGCCTCAATGAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6336
 Qy 2108 ValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsnLeuGlnLysSerLysSerAsn 2127
 Db 6337 GTGGGAAGTGGCTCTCTTCAAGATTTCAACATCAGCAATTTGAGAGAAATCCATCAGCAAC 6396
 Qy 2128 ProProGlySerAsnLeuArgThrThr 2136
 Db 6397 CCCCCAGGCTCCAACTCGGAGCACT 6423
 Db
 RESULT 15
 AR302612 6885 bp DNA linear PAT 12-JUN-2003
 LOCUS AR302612 Sequence 49 from patent US 6541252.
 DEFINITION AR302612
 ACCESSION AR302612
 VERSION AR302612.1 GI:31690910
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 6885)

AUTHORS Walke,D.W., Hilbun,E., Donoho,G. and Turner,C.A. Jr.
 TITLE Human kinases and polynucleotides encoding the same
 JOURNAL Patent: US 5541252-A 49 01-APR-2003;
 FEATURES Location/Qualifiers
 source 1..6885
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 3,41e-243 Length: 6885
 Score: 10211.00 Matches: 2048
 Percent Similarity: 88.20% Conservative: 0
 Best Local Similarity: 88.20% Mismatches: 0
 Query Match: 94.44% Indels: 274
 DB: 6 Gaps: 2

US-10-010-720-14 (1-2136) x AR302612 (1-6885)

QY 61 MetAspLysAspSerArgGlyAlaAlaThrThrThrThrGluHisArgPhe 80
 Db 1 ATGGACAAAGACACCGCGTGGGGCGCGCGACACCACTACCACTGAGCACCCTTCTT 60
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluMetProGlyLeuProLeu 100
 Db 61 CGCGGAGCGTCACTGCGACTCCATGCACTGCACTGAGAGCTTCCGACCTTCTCTT 120
 QY 101 SerLeuProGluProSerIleProAlaAlaValProGluSerAlaProProGluProHis 120
 Db 121 TCCCTGCCCCAGCCAGCATCCCGCGGCTGTCGCCAGAGAGTCTCACGAGCCCCAC 180
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGluValAlaGluGluProProAlaAla 140
 Db 181 CGGGAAGACACCGTGAACCGCCACCGCACTTCCAGTAAAGCCAGAGCTCCAGCGCT 240
 QY 141 AlaAlaProGlyGluGluAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 241 GCCGCCCTGGGGAACAGGCCCTCGCGGCCCTTCCCTCCACTGTCCTCCAGCAGTACC 300
 QY 161 SerLysAspArgProValSerGluProSerLeuValGlySerLysGluGluProProPro 180
 Db 301 AGCAAAAGACCGCCAGTGTCCAGCCTTGTGGGAGCAAGAGAGAGCCGCCCGC 360
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGluGluArgSerGluGln 200
 Db 361 GCGAAGATGGGAGCGCGCGCGCCAGCCCAAGAGCCACAGAGAGAGCCAGCCAGCAG 420
 QY 201 GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
 Db 421 CAGATGATATCGAAGAGCTGAGACCAAGGCCGCGGAAATGCTTAACGATGGCCGCTT 480
 QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
 Db 481 CTCAAGTTTGACATCGAAATCGGACAGAGGCTCTTTAAGCGCTCAAGAGCTGAC 540
 QY 241 ThrGluThrThrValGluValAlaTyrCysGluLeuGluAspArgLysLeuThrLysSer 260
 Db 541 ACTAAACACCGCGGAGAGTGCCTGCTGTAACCTGACAGATCCAAAATTAAACAAGCT 600
 QY 261 GlnArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGluHisProAsnIle 280
 Db 601 GAGAGCGAGAGATTAAAGAAAGAGCTGAATGTTAAAGGTCTTCCAGCATCCCAATATT 660
 QY 281 ValArgPheLysAspSerTyrGluSerThrValLysGlyLysLysCysIleValLeuVal 300
 Db 661 GTTAGATTATATATCTCTGGAATCCACAGTAAAGAAAGAAAGATGCTTTTGGT 720
 QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
 Db 721 ACTGAACCTTGAGCTTGGAACAATTAAACGATCTGAAAAGGTTTAAAGATGAGAG 780
 QY 321 IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGluPheLeuHisThr 340

Db 781 ATCAAACTCTAAGAGCTGTGCGCTCAGATCCTTAAAGTCTTCACTTCTTCACT 840
 QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 Db 841 CGAAGCTCACTATCATTCACCCGATCTTAAATGTGACAAACATCTTATCACCGGCCCT 900
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysAlaGluAspPheAla 380
 Db 901 ACTGGCTCACTCAAGATTTGAGACCTGCTGCAACCTTGAAGCGGCTTCTTTGGC 960
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
 Db 961 AAGAGTGATAGTATGCCCAAGTTCATGCCCCCTGAGATGTATGAGGAAATATGAT 1020
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
 Db 1021 GAATCCGTTGACGTTATGCTTTGGAGTGTGATGCTTGAAGTGCATCTGAATAT 1080
 QY 421 ProTyrSerGluCysGluAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
 Db 1081 CCTTACTCGAGTGCCTCAAAATCTGACAGATCTTCCGCTGACCAAGTGGGATGAG 1140
 QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleGluGluCysIle 460
 Db 1141 CCAGCCAGTTTGAACAAAGTACCAATCTCTGAAGTAAAGAAATTTTGAAGATGCATA 1200
 QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
 Db 1201 CGACAAAACAAAGAAAGATATCCATCAAAAGACCTTTTGAACATGCTTCTTCCA 1260
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
 Db 1261 GAGGAAACAGAGTACGGGTGAATTAACAGAAAGATGTGGAAGAAATAATACCATTA 1320
 QY 501 LysLeuTyrLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
 Db 1321 AATTATGCTACGATATTGAAGATTTTAAAGGAAATTAACAAATATATAGAA 1380
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet 540
 Db 1381 GCTATTGATTTTCTTTTGAATTTAGAGAGATGCTCCAGAAAGTGTGCACAGAAATG 1440
 QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 Db 1441 GTAAAGTGTGGATGATGTGAAGGTGATCAACAGACCATGGCTTAAAGTATCAAGAC 1500
 QY 561 ArgValSerLeuIleLysArgLysArgGluGluArgGluLeuValArgGluGluGln 580
 Db 1501 AGAGTATCATTAATTAAAGAGAAACGAGACGCGCATGTTGTAAGGAGAGCAAGAA 1560
 QY 581 LysLysLysGluGluGluSerSerLeuLysGlnIleValGluGlnSerSerAlaSerGln 600
 Db 1561 AAAAAAAGAGAGAGAGACAGTCAACAGACAGTGAACAATCCAGTCTTCCAG 1620
 QY 601 ThrGlyLysLeuGluProSerAlaSerThrGlyLysProThrAlaSerThrThrSer 620
 Db 1621 ACHAGATCAAGAGCTCTCTTCTGAGACCGGCACTACTCTGCTTCACTCACTTCA 1680
 QY 621 AlaSerValSerThrGluValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
 Db 1681 GCTTCAAGTTCTACACAGTAGAACCTGAAGAACCTGAGCAATCAACATCAACAATA 1740
 QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 Db 1741 CAGTACCAAGCAACCGATATCTGTATTATGATGAGCGGCTTGAACATGTGACAGGA 1800
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGluThrValSerTyrGlySerGln 680
 Db 1801 TCTCTGCTTCAACAATCTCAGTGAAGCAAGCAAGCAATTTCAATATGTTCCCA 1860
 QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 Db 1861 CATGAACAGGACATCTACAGGCAAGTCCAGGCAATATCTTACTGTCCAGCA 1920

QY	701	GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnSerGlnGly	720	QY	815	ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer	834
DB	1921	CAGCTCAGGCCCATGGGATATATCCACCTCAAGTGTG-----	1959	DB	2917	CCGACCACTTTGGCTTCTCTGTAGACATGCATTCAGATGTTGCTTCAGGATATAGT	2976
QY	721	GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740	QY	835	AspGlyAsnGlnLeuValProSerSerSerGlyArgHisGlnGlyValGlyThrThrLeuArg	854
DB	1959	-----	1959	DB	2977	GATGGCAATGAGAACGCCATCTTCAGTGGAGGAGGATAGAGAACTACCAAAACGG	3036
QY	741	GlnGlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer	760	QY	855	HisTyrArgIleSerSerValArgSerArgSerArgHisGlnIleThrSerArgProIleLeu	874
DB	1960	---CAGCAGGAGAAATACAGCAGACAGCCCTCTCCACAGCAGAGTATTCACCTTCA	2016	DB	3037	CATTACCAAAATCTGTAGAGAGCGCTCTCGACATGAAAAAATTACCGCCCAAAATTA	3096
QY	761	GlnThrSerThrSerSerSerGlyAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780	QY	875	ArgIleLeuAsnValSerAsnIleGlyAspArgValAlaGlnCysGlnIleGlnThrHis	894
DB	2017	CAGACATCAACCTCCAGTAGGAGCCACTACTGCACAGCCAGTAGAGTCAAGCTCAAGCTCA	2076	DB	3097	AGAAATTTGATGATTTCAATTAAGAGACCGAGTATGATGATATGCAATTAGAGACTCAT	3156
QY	781	GlnValLeuProGlnIleValSerAlaGlyIleGln-----	791	QY	895	AsnArgIleMetValThrPheIlePheAspLeuAspGlyAspAsnProGlnIleIleAla	914
DB	2077	CAAGTCTTGCTCAAGTATCAAGCTGGAAACAGCTTCCAGTTCCAGCCAGTACCAACT	2136	DB	3157	AATAGAAATAGTTTACATTCAAATTTGACTTAAGTGTGCACAAACCCGAGAGATAGCA	3216
QY	791	-----	791	QY	915	ThrIleMetValAsnAsnAspPheIleValAlaIleGlnIleArgGlnSerPheValAspGln	934
DB	2137	ATCCAAGGCGAACTCAGATCCAGTTGCGACACAAACCTCGTGTGTTCCAGTCCACTCT	2196	DB	3217	ACAAATTATGGTGAACATGACTTTATTTAGCAATAGAGAGAGAGTGGTTGTGATCA	3276
QY	791	-----	791	QY	935	ValArgGlnIleIleGlnIleValAlaAspGlnMetLeuSerGlnAspValSerValGlnPro	954
DB	2197	GGTGTCTATTTCTTCCAGTAGGAGACGCCCTCCTACTCCTCTGCTCCCTCAGTACCT	2256	DB	3277	GTGGAGAAATTAATGAAAAAGCTGATGAATGCTCACTGAGATGATGATGATGATGATCA	3336
QY	791	-----	791	QY	955	GlnGlyAspGlnIleGlnIleGlnSerIleGlnIleGlnIleGlnIleGlnIleGlnIleGln	974
DB	2257	GTTCTCTCAGATTCCCATATCAATCTCTCATGTGTCTACGGCTCAGACAGGTTTCTCATCC	2316	DB	3337	GAGGTGATCAGGGATGAGAGTGTACAGAGAAAGATGACTATGCTTTTCAGGTTCT	3396
QY	791	-----	791	QY	975	GlnIleLeuGlnGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGln	994
DB	2317	CTTCCCATCACAATGGCAGCTGGCATTTACTCAGCTCTGCTCAGCTGGCTTCATCTGCT	2376	DB	3397	CAGAAATTTGAGAGAGATTTCAACAAACAAATCTCGGCTCTTCATGCCACAGCAAAATA	3456
QY	791	-----	791	DB	995	GlyIleProThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgPheIleVal	1014
DB	2377	ACAAAGCTGCGATCCCGGGGATATCACTGTGTCTAGTCAAGTTCMAACCTCTTG	2436	DB	3457	GGCATTCCTACAGATCTTTTAACTCAAGTGTTCATTTGCGCGAAGGGGTTTATAGT	3516
QY	791	-----	791	QY	1015	SerProValProGlnSerArgLeuArgIleGlnSerIleValPheProSerGlnIleThrAsp	1034
DB	2437	CAGCTGTGACTCAGCTGCCAAGTCAAGTTCAACCCACAGCTCTCTACAAACAGAGTTGAG	2496	DB	3517	AGTCTGTGCGCAAGAGCCGATTTACGAGATCAAAAGTTTCCCGAGTAAATTAACAGAT	3576
QY	791	-----	791	QY	1035	ThrValAlaIleSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer	1054
DB	2497	TCCATGGGAATACAGCTAACTTGGACAAGCTGTAGAGTTCCACTTCTCTGGAGAT	2556	DB	3577	ACAGTTCCTGCTCTCAACACTCAGAGCCCTGGAATGAATCTGTCTCATCTGCATCATCC	3636
QY	791	-----	791	QY	1055	LeuSerLeuGlnGlnAlaPheSerGlnLeuAlaArgAlaGlnMetThrGlnIleGlyProAsn	1074
DB	2557	GTTCTGTACAGGCTTCCCACTCGACTGCACTGCCACACAGATACCCAGAGATTCAAATAT	2616	DB	3637	CTTAGCTTACCAACAGGCTTTTCTGAACTTACAGCTGCGCCCAATGACAGAGAGACCAAT	3696
QY	791	-----	791	QY	1075	ThrAlaProProAsnPheSerHisThrGlyProThrPheProValAlaProProPheLeu	1094
DB	2617	GCTCCCTCTTCAAGTGGCTTCTGTGTGATCATTTCTACAGTCTCTAMCCCTCCCATG	2676	DB	3697	ACAGCACTTCCAACTTTTGTCTATACAGACCAACATTTCCAGTATACCTCTTCTTCA	3756
QY	791	-----	791	QY	1095	SerSerIleAlaGlyValProThrThrAlaAlaIleThrAlaProValProAlaThrSer	1114
DB	2677	CCGACAGAACTACTGGCTACACTGGGTACTTCCACAGTGGTGACGCTTATATGGAA	2736	DB	3757	AGTAGATGCTGGAGTCCCAACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	3816
QY	791	-----	791	QY	1115	SerProProAsnAspIleSerThrSerValIleGlnSerGlnValThrValProThrGln	1134
DB	2737	TCAAACTTTTATTTCTATGGGTGTGTAGAGAGACAGTTCAGTGTCCCAAGCAGGA	2796	DB	3817	AGCCCTCTTAATACATTTTCCATCATCAGATTAATTCAGTGTGATACAGTGTCCCAAT	3876
QY	792	-----	792	QY	1135	GlnGlyIleAlaGlyValAlaThrSerThrGlyValValAlaThrSerGlyIleProIle	1154
DB	2797	GGGAGTTTACACAGCCCACTACATCTCTCCACAGACAGTGTGGAGAGTACTCAG	2856	DB	3877	GAGGGATGCTGAGTGTCCACACAGCAGAGTGTGTATCTCAGGTGTCTCCCATATA	3936
QY	795	GlyValSerGlnValAlaProAlaGlnProValAlaValAlaGlnProGlnAlaThrGln	814	QY	1155	ProProValSerGlnSerProValIleSerSerValIleSerSerIleThrIleProAla	1174
DB	2857	GGAGTCTCTAGGTTGCTCTCGCAGAGACAGTTCAGTAGACAGCCCAAGCTACCCAG	2916	DB	3937	CCACTGTGTCTGATACCAAGTATCTTCAGGCTAGTTCAGATTCACAAATACCTGCA	3996
QY				QY	1175	ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGlnIle	1194

Db 3997 GTTGTCTCAATATCTACTACATCCCGCTCATTCAAGTCCCCACATCCACATCTGAGATC 4056
Qy 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAla 1214
Db 4057 GTTCTTTCTGTAGTACAGACTGTATCTTCAGTACAGTTTCAAGCACTTCAAGCTTCA 4116
Qy 1215 GlyGlySerThrAlaThrProGlyProLysProProAlaValValSerGlnAlaAla 1234
Db 4117 GGGGGCAGTACTGCTCCACCGAGTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 4176
Qy 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPhePro 1254
Db 4177 GAGAGCACTACTGTGGAGCCATTTAACTCACTTCAAGCTTCAAGCTTCAAGCTTCA 4236
Qy 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrThrProThrIleAlaGln 1274
Db 4237 ACAGCTTCAAGCTGCTGCACTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 1294
Qy 1275 ThrValValValSerAlaHisSerLeuAlaPheSerThrSerThrThrGlyLeu 1314
Db 4297 ACCGGTAGTATTAGGCACTCACTCAATGATTAAGATCTCAATAGCAAGTCAACTGGATTG 1334
Qy 1295 AlaPheSerLeuSerAlaProSerSerSerSerProGlyAlaGlyValSerSerTyr 1354
Db 4357 GCTTTCTCCCTCTCTGACCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1374
Qy 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1394
Db 4417 ATTCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1414
Qy 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValIleProSer 1434
Db 4477 ATTCTTCCCAAGCAGACAGCACTTCTTCACTTATTTACCCCAAGTCACTTATTTAC 1456
Qy 1355 ProProLeuValGlnProValAlaAlaValProAlaValGlnThrLeuLeuHisSer 1474
Db 4537 CCACCTTGGTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1494
Qy 1375 GlnProGlnProAlaLeuLeuProAlaGlnProHisSerHisSerProGlyValIlePro 1514
Db 4597 CAGCTTCAACCAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1534
Qy 1395 AspThrGlnProLysAlaProGlyIleLeuAspIleLeuGlnGlnGlnGlnGlnGlnGln 1554
Db 4657 GATACACACCAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1574
Qy 1415 SerLeuPheSerGlnHisSerSerSerSerGlyAlaGlnHisSerValSerLeuGlnThr 1594
Db 4717 TCTCTGTTCAGTACACAGCTATCTGAGCTTCAAGCTTCTTCTTCTTCTTCTTCTTCTT 1614
Qy 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1634
Db 4777 TCACTGTATTAAGAGCACTGTCACACAGGATCCCACTTCTGTTGACCAAGC 1654
Qy 1455 LysLeuLeuThrSerThrThrSerThrGlyLeuProProThrAsnLeuProLeuGlnThr 1674
Db 4837 AAACCTCTACTTCTTCAACCAAGTACTTCTTCAACCAAGTACTTCTTCAACCAAGT 1694
Qy 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1714
Db 4897 GTTGTCTTGGCCAGTACACAGCTGCTCAAGCTTGGGCAAGTTTCAACCCAGTCAAGCT 1734
Qy 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1754
Db 4957 ACTACATAGAGAGGAACTGTAAGTCTCTCTCTCAAGCCCACTTAACTAAAGTCTCGC 1774
Qy 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1794
Db 5017 GTGCTGCAGAGGATCTGAACTTCTGACAGGTACTTCAACCAAGCAGCAGCTGCAAGCT 1814
Qy 1535 PheProGlyProSerLeuThrGlnSerGlnProLeuGlnAspLeuAlaGlnLeu 1834
Db 1554 TTTCCAGAGCCTTCTTCAACCAAGTCCAGCACTCTAAGAGATCTTGATGCTCAAT 1854
Qy 1555 ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValGlyProValSerMet 1874
Db 5137 AGAAGAACCTTGTAGTCCAGAGATATCAAGTACTTCTGCGTGTGCTCTGCTGCTGCTGCT 1894
Qy 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnGlyValSerGlnVal 1914
Db 5197 GCGGCTCCAAAGCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1934
Qy 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1954
Db 5257 AAGAGAGGCTCTGCTCTCAACCAAGTCTTCAAGAGCTGCTGCTTCTTAAAGTGGAGCATTT 1974
Qy 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAla 1994
Db 5317 CAGGTTCTGTTGCGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2014
Qy 1635 LysSerValHisPheGlnSerSerThrSerGlnSerValLeuSerSerSerPro 2034
Db 5377 AAGTCTGTTCAATTTGAATCCAGCACTCAAGATCTTCAAGTCAATCAAGTCAATCAAGT 2054
Qy 1655 GluSerThrLeuValLysProGlnProAlaGlnLysSerValIleThrIleProGlyIleSerSerAsp 2074
Db 5437 GAGAGTACCTTGTGGAAGCAGAGCGAATGGATTAACCATCCCTGATCTCTTCAAGT 2094
Qy 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlnPro 2114
Db 5497 GTGCCAGAGAGTCCCAACCAAGTCACTGCTCAGAGGCAAGTCAACCAAGTCAACCAAGT 2134
Qy 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 2154
Db 5557 ACCAAGTGGAGCGTTTCAAGTGAACCACTAAGCAACCAAGTGGGCTGCTTCTCTGTA 2174
Qy 1715 SerLysThrGlnAspLysIleThrAspThrLysGlnGlyProValAlaSerProPro 2194
Db 5617 TCAAAACCTGAGAGACAAATCACTGACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 2214
Qy 1735 PheMetAspLeuGlnAlaValLeuProAlaValIleProLysGlnLysProGln 2234
Db 5677 TTTATGATTTGGAACAAGCTGTTCTTCTCTCTGATGATCAACAAAGAAAGAGAGAGAG 2254
Qy 1755 LeuSerGlnProSerHisLeuAsnGlyProSerSerSerProGlnAlaAlaPheLeuSer 2274
Db 5737 CTGTCAAGAGCTTCAACATTAATAGGCGCTTCTTCAACCGGAGCGCTTTTAACT 2294
Qy 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerLysSer 2314
Db 5797 AGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2334
Qy 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyrMetSer 2354
Db 5857 CTTCTCAAGCAGATTAAGTCAAGGCTTAACTTAATTAATTAATTAATTAATTAATTAAT 2374
Qy 1815 SerAspAspGlnSerAspIleGlnAspGlnSerAspLysGlnLeuArgArgLeuArg 2394
Db 5917 AGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2414
Qy 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerAspLysHisGlnIleGlnSer 2434
Db 5977 GATTAACATCTCAAGAGATTCAGAGCTGCAAGTGGCCAGAGAGAGAGAGAGAGAGAGAG 2454
Qy 1855 LeuTyrThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu 2474
Db 6037 TTGTATACCAAGCTGCGCAAGGTGCGCTGCTTATTAATTTCCCAAGCTGCTCTCCCTT 2494
Qy 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSer 2514
Db 6097 TCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2534
Qy 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 2554
Db 6157 TTGGGAAATTAAGCCCTTCAAGCTTCTGATGATGATGATGATGATGATGATGATGATGAT 2574

QY 1915 LeuHisProGlnGlnThrLeuHisProGlnGlnSerGlyGlnAsnGln 1934
 DB 6217 TTGGACCCCGAGAGACCCCTCCAGCTCTGGCAGCATCCGAGAGTCCGGGAGAGATCAG 6276
 QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuTyrSerAlaPheThrSer 1954
 DB 6277 CTGTTACAGCCCTTAAGCCATCTCCCTCAGTGACACCTCTATTCAGCCTTCAACAGT 6336
 QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
 DB 6337 GATGGTGCATTTCACTACCAAGCCTTCTGCTCCAGGTCAAGGACCAAGACAGACAAAC 6396
 QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSer 1994
 DB 6397 ACTGTTGGGCGCAAGTGAACAGCAGCCGCGCCAGCTCAGCTCCTGCGCATGAGCTCC 6456
 QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAsp 2014
 DB 6457 AGCAGGAAGGCGCATTCACAGATGACTTGCACAAAGTTGGTAGACAATTGGGCCCGAGAT 6516
 QY 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrGlnGlyProGly 2034
 DB 6517 GCCATGATCTCTCAGGAGAGAGAGAGCAAGGCGCATGATGATGAGAGGCGCTTGA 6576
 QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
 DB 6577 ATGGCAAGGAATCTCTGCGACCTGGGCAACTGTGCATCTCCATGACCTCGAACCCTGGGT 6636
 QY 2055 GlySerAlaProIleSerAlaIleSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074
 DB 6637 GGCTCTGCCCCCATCTCTGAGCATCAGCTACTCTCTAGTCACTTCAACAAAGTCTATG 6696
 QY 2075 CysProGlnGlnTyrGlyPheProAlaThrProPheGlyAlaGlnTrpSerGlyThr 2094
 DB 6697 TGGCCCCCAGCAGCATGTGGCTTCCAGTACCCCATTTGGCGCTCATATGAGTGGAGCG 6756
 QY 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
 DB 6757 GGTGGCCCGAGACAGCAGCCTTGGCCAGTTCACACCTGTGGGAACCTGCCTCTTGAG 6816
 QY 2115 AsnPheAsnIleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArg 2134
 DB 6817 AATTCAACATCAGCAATTGCGAATTCATCAGCAACCCCGAGGCTCCCACTCGCGG 6876
 QY 2135 ThrThr 2136
 DB 6877 ACCACT 6882

Search completed: September 25, 2004, 08:40:52
 Job time : 12502 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:00:53 ; Search time 64 Seconds
(without alignments)
3210.392 Million cell updates/sec

Title: US-10-010-720-14
Perfect score: 10812
Sequence: 1 MSGAAEKOSSTPGSLFLSP.....NISNLOKISINPPGSNLRRT 2136

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1656	15.3	1851	2 T19964	hypothetical prote
2	765.5	7.1	633	1 S49611	probable serine/th
3	745.5	6.9	516	2 T06716	hypothetical prote
4	702.5	6.5	547	2 T46059	MAP kinase - Arabi
5	702.5	6.5	549	2 T51039	MAP kinase (import
6	678.5	6.3	677	2 T02951	probable mitogen a
7	591.5	5.5	2232	2 T34434	hypothetical prote
8	494	4.6	3570	2 T45025	mucin MUC5B, trach
9	478	4.4	1802	2 S69703	HK1 protein precu
10	475.5	4.4	2090	2 S26058	probable transform
11	474.5	4.4	1952	2 T48814	hypothetical prote
12	472	4.4	2187	2 T30826	nascent polypeptid
13	468	4.3	1192	2 T18611	probable serine/th
14	467.5	4.3	3507	2 T34513	hypothetical prote
15	460	4.3	1367	2 S48478	glucan 1,4-alpha-g
16	457	4.2	1398	2 T13741	hypothetical prote
17	454.5	4.2	1246	2 G89287	hypothetical prote
18	434.5	4.0	2271	2 F90073	protein H39E23.1
19	431.5	4.0	792	2 JC7122	hypothetical prote
20	428.5	4.0	1777	2 E95206	hypothetical prote
21	428	4.0	4776	2 E95206	cell wall surface
22	421.5	3.9	13288	2 T03099	mucin, submaxillar
23	416.5	3.9	1630	2 A53577	ascites staloglyco
24	416	3.8	1080	2 S48944	hypothetical prote
25	416	3.8	5327	2 T13564	microtubule-associ
26	413	3.8	1459	2 T32271	hypothetical prote
27	406	3.7	1829	2 T24583	hypothetical prote
28	403	3.7	2529	2 A56923	transcription fact
29	398	3.7	1233	2 T30989	serine/threonine p

30	396.5	3.7	528	2 I47141	gastric mucin (c10
31	396.5	3.7	1734	2 A54602	microtubule-associ
32	393.5	3.6	2578	2 A56922	transcription fact
33	388.5	3.6	865	2 A47282	calcium-binding pr
34	387.5	3.6	1306	2 S25370	MSB2 protein - Yea
35	385.5	3.6	2464	1 QRMSP1	microtubule-associ
36	385	3.6	841	1 I78885	serine/threonine-s
37	383	3.5	2264	2 A56577	microtubule-associ
38	383	3.5	2453	2 S60254	nuclear receptor c
39	382.5	3.5	1609	2 S25345	probable membrane
40	381.5	3.5	1611	2 T33836	hypothetical prote
41	380	3.5	1275	2 T33369	hypothetical prote
42	378	3.5	939	2 S28394	probable serine/th
43	374	3.5	1206	2 T34021	protein kinase SK2
44	373.5	3.5	873	2 A47283	calphorin - fruit
45	373.5	3.5	1032	2 T34433	hypothetical prote

ALIGNMENTS

RESULT 1

T19964 hypothetical protein C46C2.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T19964

R/McMurray, A. submitted to the EMBL Data Library, December 1995

A/Reference number: Z19204

A/Accession: T19964

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1851 <W1>

A/Cross-references: EMBL:Z68296; P1DN:CAA92591.1; GSPDB:GN00022; CESP:C46C2.1

A/Experimental source: clone C46C2

C/genetics:

A/Gen: CESP:C46C2.1

A/Map position: 4

A/Introns: 15/3; 42/1; 65/2; 93/3; 239/2; 340/1; 448/3; 863/3; 1096/3; 1351/3; 1495/3; 1

Query Match	15.3%	Score 1656; DB 2; Length 1851;
Best Local Similarity	27.1%	Pred. No. 1.5e-41;
Matches 605; Conservative 272; Mismatches 646; Indels 710; Gaps 87;		
QY	8	KQSTPGSLFLSPAPAPKXGSSDSVC-----EKLGAAADAVTGR--50
DB	14	RMGITWGG--RPPAPSSVSTTASTGTFGRRLVNRIRKYVDLHPAQENPTWGSW 70
QY	51	-TEYRRRRRTMDKSGAATTTTEHR-----FFRSYICDNATALEPGL--PLSL 102
DB	71	LSBERSRLEAVQDMKRTPKMTSKKRPDDPTSSPSVLSNALENSTPSINNVSSI 130
QY	103	PGSPISPAVPSAPPEPHEEYVATA---TSQVAQOP-----138
DB	131	TNSSSPFSLGAA-----TSTASAIPTSVATNPHLNHVSRIPOAIVTGTING 182
QY	139	-----AAAPGEQAVAP-APST-----155
DB	183	SLPPLISPTSAAMATPLISGKAGPMSPSTPINVAATVQNAVSSPQHSIFDRSLNK 242
QY	156	VPSSTG-----KDRVSGPSLVGSK-----EEPPPARSG-----SGG 187
DB	243	IPPNSTLSSSSPSDANNDKFIQQRHSILSNVRLTQAMVNDGRTLTGDMDMQVBE 302
QY	168	GSAPKPEPERSQOQD-----DIE---ELDTKAVGKNSDGRFLKFDLEIGRSGFKVYKSL 239
DB	303	EPARKQEKREBEHGAARIVDEDFDAQEKPIDSKNGRFLKFPDEIGRSGFKVFRGL 362
QY	240	DTETTEVAWCELQORKLTKSRQPKFAEMLKGIQHNIYRFDSWST--VKGR-KCI 297
DB	363	DIETGVAWMCLEQSKINKTERQRFREAEMLKDIQHNIYRFDIYWESADLCGRKTY 422

QY	298	VIVTEMTSGTLKTYIKRFRVMKIK-VLRSWCROIILGLOFLHTRPPIIHRDLKONIF	356
Db	423	VVLTEMTSGTLKMYLKRFRINIKVVLKSWCROILKGLSFLHRNPVVIHRDLKONIF	482
QY	357	IIGPTSVYKIGDLGATTLKRASFASVYGTPEPFAPEMYEERYDESVDVYAFGMCMLEMA	416
Db	483	IGTTGSVYKIGDLGATLKNNSFAFSVGTPEPFAPEMYEEMVDESVDVYAFGMCLLEW	542
QY	417	TSEYPSFCQNAQIYRVRTSGVSKFASFDKV-AIPEYKEIIEECIQANDERSIYDOL	474
Db	543	TSEYPSFCMNPATIRYKVISGVKRECSRIIPAOYPEIRLIDCIKVRREESYVQOLL	602
QY	475	NHAFQOE--TGVRYLEAED---DGEYIAIKMLRIEDIKKLKQ-KYKQNEAIEFSFDL	528
Db	603	VDDEFTEPDLIGIRFELIKNRADLINDNVELOMOLRYAYDEKKRQYAFKNEGLQAFPI	662
QY	529	ERVDVEDVAQEMVESGYVCEGDHMTMAKAIKDRSL-----IKRERQOLVRE	577
Db	663	ENDSPDEVYQCMIBQHHPBEDTREMITKLKKQYDAFRDRDRHLEIKRAKEERERIE	722
QY	578	EQE-----KKQOESSLKQOYEGSSASQGTIKOLPSASTGIPASTTSVSVQ	626
Db	723	EALKEBELRLRAEKEREKERLEKERLEKRAAAAAANPNP---TPIPTPATPHS-SAQ	778
QY	627	VEPEP---EADHQOOLQYQPSISVLSLNGTVDSQGSVFTESRVSSQCTSVYSGQHE	682
Db	779	QQPIPEPLSTQTAHEIDQSAQQPSVFP-----	805
QY	683	QAHSTGVPOHIPSTVOAQSQPHGVYPPSSVA-----QQSQSQQPSSSLT	728
Db	806	-----TMANIPAMSPTSAQPOQPVLPISAAVNPPTMIHVPKSELPVQNVATTAFAV	859
QY	729	GVSSSQPIHQPOQOQGIQCTAPRQCVQVYSLQTSSTSEATTACQVSOAPQVLPQVSA	788
Db	860	AAANNVPSPAPFKTEIDIQPTLAQMTVPRTISTDSGLVINTPASIAF-----SP	910
QY	789	GKQSTQSVQVAPAEVAVAAQOAPQPTTLASVDSAISDVASGMSGDENVPSSGRHE	848
Db	911	APSATDVASTTAPVP-----ALPPTT-----TTQGAAMAASTTENK	948
QY	849	GRTYKRYKYSRKSREKTSRPKRLILNV-SNKGDVEVCOLETHNRKRYTEKPELDG	907
Db	949	EERKSNKRYV-----MEITGGDESNFALVSCRIDT-SHKSVTFFQEPAGT	994
QY	908	DNPEIATIMVNDPFLIAERESFVDQREILIEADMLSEDSVEBEGDQLESIQKD	967
Db	995	DKPCTIAIKTLAEDOLKVAHVHIVEAOGEVY-----OLINSD-----GKK	1035
QY	968	DYFSGSQKLEGEFKOPIRPASS-WPOOLGRTSLSLQVHASAGRFLVSPVPSRLRESK	1026
Db	1036	GVGTGLATVLDPNSTEPTPIIAVWPKXSSAATASNTK-----PKIEIKTEPPT-RDAS	1087
QY	1027	VFPEIIDTVAASTAQSBGMNLSHSASLSLQAFSELRAPQMTGEPN-----TAPNPJAH	1082
Db	1088	QEFNNVOVTVKRYSOE-----SNABSVQSIIRPPGGIIVMSPTNQTD	1129
QY	1083	TGPTFPVVPFLSIAGVPTTAATAAPVATSSPNDISTSVIQSEVTVPTBEGIAGVAT	1142
Db	1130	SAP-----PPTGAAPD-----	1141
QY	1143	STGVTVSGCLIPPVSESVLSSVWSITPIPAVYSITSPSLQVPTS-TSRIIVVSTAL	1201
Db	1142	SRFOVTVKSADPI-----ATPISSSISITATVPIVAA-----TPNITSSPVI-	1183
QY	1202	YPSVTVASATSASAGSTATPGPRPAPVNSQQAAGSTTVGATLTSVITTSFESTASQLSI	1261
Db	1184	-----VQPITAQVITHLAPPSVSHS-LSSNSPSKATTHSNMSIQSTTSP-----	1229
QY	1262	QLSSSSTSTPLAETVAVSAHSLDKTSHSSTGLASLSAPSSSSSPGAGVSVYSIQPGL	1321
Db	1230	-----GRRPTVQPVQQAES-----GLSSSISTFP-----	1252
QY	1322	HHVLVPSVIASTPILPQAAQPTSTPLDLQVPSIPIPLVQVPAVNVPAVQOQTLHNSPQAL	1381

```

Db      1253 HPEPTAI-----TSCP--PVPSVPPV-----VSNGLNLEVPKQT 1288
QY      1382 PNOPIHCEVSDTQPKACGIDDIKTLIEKLSLFSHSSGGAOMASVLSLETSLVBEST 1441
Db      1289 PSATNMNV-----DTC-----HSSSTASTALVSETPATVH-- 1319
QY      1442 VTP-GIPTAAVAP-----SKLLTSTSTCLPPTNPLGLVVALPVTPVTPQOVSTPVST 1494
Db      1320 VTFISVPAVQEPVLVDHHSDDLTC-----LDSELKRVSGVSHSASPSTVSES 1367
QY      1495 TTSGVKPGTAP-----SKPELTAPVLPVGTETLPAGLTSPQLEPFPFQ-----BELTOSQ 1544
Db      1368 LTS--MTPOITPLACQTFVPASIGQPAV--IAAIAIASIIPNASVQSPSLDAETGLAGIH 1425
QY      1545 QPLEDL--DAQURLSPKXITVTSAVG--FVSNAAPTA----- 1579
Db      1426 EKLEAKKEQDRREDMDGDAITTTTDDKDEIPIIDLKGLAALGKVIHMDGRETTMP 1485
QY      1580 ----ITNAGTOPQOKGVSVQVKEGPVLATSSGAGVGNKGRFQVSVAAAGAOKKKNKSEDAK 1635
Db      1486 DHPDLTASTQ--LISPSNPDVLITMSSA-----VEGASSTWIE 1524
QY      1636 SVHEESTSESVLSSSSPESSTLYKE--PNCITIPGISSDVPEPSAKHTTASEAKSDTGP 1634
Db      1525 DIASTSAVAVASMMNSMPPGAQNSTDQIPAAVT--LSMD-QECAQSMTSSITRNTTG-- 1578
QY      1695 TKYGRFVUTTANKVGRFVSFKTEDKITDTYKKGAVASPPFMDLEQAVLPAVTPKKEKPE 1754
Db      1579 TKLATPF-----NLFTALSTLTGTHIRQP- 1602
QY      1755 LSEPSHLNGSSDPEA-----AFLSRDVDGSGS-----PSPHQLSSKLPS 1797
Db      1603 ----NAPSSRDETAPMTPISTNERIGGGGGGAGATFSIGTPPSHP-----FPV 1648
QY      1798 QNLSQSLNSFNSSTYSSDNESDIEDLKLRLRLRDKHLKEIODLSROKHEISTLYT 1857
Db      1649 SECVDYDK-----QOMDLESED--PEVIQMIVRHMEGKHLKEQKVELERLRS 1695
QY      1859 KLGKVPAYIIPAAFLSGRRRPTKSGKSSKSSSLGNKSPQLSGNLSGSAASVLAHP 1917
Db      1696 KI-RVPRAISVMP--EMIGDDEADT-----LTLTALQSLAGNASLSLPAS----- 1736
QY      1918 QOTLHPNGNIPESGONOLQPLKPSPSDNLYSAFTDGAISVSLAPQCGTSTNTYVG 1977
Db      1737 ----PPTETTKVNTVITP-----SDVLAIRMTM-----SQSSTKSNVS 1773
QY      1978 ATVNSQAQAQAP 1990
Db      1774 VSSRRHONQSAAP 1786

```

RESULT 2
 S49611
 probable serine/threonine-specific protein kinase pkpa (EC 2.7.1.-) - *Phycomyces blakesleei*
 C:Species: *Phycomyces blakesleei*
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S49611; S59578
 R:Ruiz-Perez, V.; Murillo, F.; Torres-Martinez, S.
 Submitted to the EMBL Data Library, November 1994
 A:Description: A novel serine/threonine protein kinase in the fungus *Phycomyces blakesleei*
 A:Reference number: S49611
 A:Accession: S49611
 A:Molecule type: DNA
 A:Residues: 1-633 <RU1>
 A:Cross-references: EMBL:Z46636
 R:Ruiz-Perez, V.L.; Murillo, F.J.; Torres-Martinez, S.
 Curr. Genet. 28, 309-316, 1995
 A:Title: PKpa, a novel *Phycomyces blakesleei* serine/threonine protein kinase.
 A:Reference number: S59578; MUID:96120859; PMID:8590476
 A:Accession: S59578
 A:Molecule type: DNA
 A:Residues: 1-9;39-346 <RUW>

A:Cross-references: EMBL:Z46636
 A:Note: only part of the coding region is given
 C:Genetics:
 A:Gene: PKPA
 A:Introns: 60/1; 248/3; 277/2; 339/3; 534/1
 C:Superfamily: Phycomyces blakeleanus probable serine/threonine-specific protein kinase
 C:Keywords: phosphotransferase, protein kinase
 F:34-289/Domain: protein kinase homology <KIN>

Query Match 7.1%; Score 765.5; DB 1; Length 633;
 Best Local Similarity 35.2%; Pred. No. 5.5e-16;
 Matches 192; Conservative 91; Mismatches 185; Indels 77; Gaps 22;

```

QY 187 GGSAAKEQERSQOD--DIELETKAVGMSNDGRFLKFDIEIGSGFKTVYKGLDTETT 244
DB 5 GGVMNNDMEFVDOSNPDYE---KYLEASGNRGYSKLTNLVLKGAIVKYKAIIDBEA 60
QY 245 VEVAMCELODRKLT--ESERQREKE---EAMMLGLQHPNIVRYFYDSWESIVKAKCIVL 299
DB 61 I-----NDNEITNVAVTRQEFKDLGHEIDILKSVPHPNITPHDAMYETE----FVF 109
QY 300 VTEIMTSGTLKTYLKRFRKM-KIKVLRSGCQILKGLQFLHTRPTPIIHRDLKCDNIFIT 358
DB 110 ITLEMTSGTLREYIRKLTPLPNIKIYKRCQILKGLAYLHGHPPIIHRDLKCDNIFIN 169
QY 359 GPTGSAVIGDLGLATLKRAFPKSVITGTPFMAPEMEYEEK-YDSVDVYAFGMCLEMAT 417
DB 170 GAHEIKIGDMGTAEMNGK-KYTVIGTPFMAPEMEYEEKVDIVAFGMCLEMAT 228
QY 418 SEVYSESCQNAAGIYRVTSVGRKAFDVAIPEVKEIIIGCIRONDERSYIKDLINHA 477
DB 229 GEYFPGCTYAVGVFKKVTOTIKPECSRVQDEPLTLVNI CLTP-EDEMTQAEILEHR 287
QY 478 FFOETGVRVELAEEDGEEKIAIKMLRIDIKKAKYKONDAIEFSPFLERDVEDVA 537
DB 288 FLAVER--EYVLVSKMTML---LTLQV---VFKG--MDKLSVKEFPAADDTTAADV 336
QY 538 QENVESGYV-----CE-----GDHKTMAKIKDVSILIKKRE 570
DB 337 AEMIEQVLANCYQOILITCEINRLDIANQGPDPKGEDEKIVMRENDIRSELBAKK 396
QY 571 QROLVRE--EOKKKKQ--EESLKKQVBSASQICIKLPASGICPIATTSASVST 625
DB 397 DLALAVRFEAKKCELEKHNI---IAERCKET-IFALEQAKQIDPLLOPOPQOP 452
QY 626 QVEPE--EPEADQHQLOQYQPSISVLSDGTVDQSGSSVFESHVSQQTYSQGH 682
DB 453 QPQPOQPOQPOQLOQLOQLSPOSTTSRPTSDSDNSTSTMLS--LSESLKLCVSGDE 511
QY 683 QAHST 687
DB 512 QVETT 516

```

RESULT 3

T06716
 hypothetical protein T29H11.220 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #ext_change 17-Mar-2000
 C:Accession: T06716
 R:Querier: F.; Choise, N.; Robert, C.; Brothier, P.; Wincker, P.; Cactolico, L.; Artigu
 submitted to the Protein Sequence Database, April 1999
 A:Reference number: Z15793
 A:Accession: T06716
 A:Molecule type: DNA
 A:Residues: 1-516 <QUB>
 A:Cross-references: EMBL:AL049659; GSPDB:GN00061; ATSP:T29H11.220
 A:Experimental source: cultivar Columbia; BAC clone T29H11
 C:Genetics:
 A:Gene: ATSP:T29H11.220
 A:Map position: 3
 A:Introns: 23/3; 36/2; 112/2; 187/1; 239/3; 295/1; 338/1
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

Query Match 6.9%; Score 745.5; DB 2; Length 516;
 Best Local Similarity 41.9%; Pred. No. 1.7e-15;
 Matches 169; Conservative 67; Mismatches 112; Indels 55; Gaps 12;

```

QY 194 QERSQOQDDIELETKAVGMSNDGRFLKFDIEIGSGFKTVYKGLDTETTVEAVM--CE 251
DB 3 QDENNSSEEFVE-----IDPTGRGRYKKEVIGKAFKEVYAFQDLSEIVAMNQVK 54
QY 252 LQDRKLTKESEKQPEAEMLKGLQHPNIVRYFYDSW---ESTVKGKKCIYLVTEIMTSG 307
DB 55 LDDKFCSSEDLRLYSVHLKTLKHSIKIFYSMIDHQMTIN-----LIEVFTSG 108
QY 308 TLKTYLKRFRKMKIYLRSGCQILKGLQFLHTRPTPIIHRDLKCDNIFITGPGSYKIG 367
DB 109 NLRYRKHKCVDIRALKKWSRQTLSEGLVYLHSDPVIHRDLKCDNIFITNGQGEYKIG 168
QY 368 DLGLAT-LKRAFPKSVITGTPFMAPEMEYEEKYDSVDVYAFGMCLEMATSEVPSECC 426
DB 169 DLGLAATLHRASASHVITGTPFMAPPELYEDYNVLVDIYAFGMCLELVTFFPYSECT 228
QY 427 NAAQIYRVTSVGRKAFDVAIPEVKEIIEGCIKONDERYSIKDLINHA--FOET 483
DB 229 NAAQIYRVTSVGRKAFDVAIPEVKEIIEGCIKONDERYSIKDLINHA--FOET 287
QY 484 -----GVRELAEDDG-----EKIAIKMLRIDIKKAKYKONDAIEFSPFLERDVEDVA 520
DB 288 ENVSHKENGNGNGVYDKLSDSEVGLTVYEGGRKDLNTIFLRLTID--SKQIRN-- 342
QY 521 AIEFSPFLERDVEDVAQAEVWS-GYVCEGDHKTMAKIKDVSILIKKRE 562
DB 343 -IHFPNIEITDSFSAIENVEBELDLDDODISTIAKMDTEI 384

```

RESULT 4

T46059
 MAP kinase - Arabidopsis thaliana
 N:Alternate names: protein T18N14.10
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #ext_change 04-Feb-2000
 C:Accession: T46059
 R:Querier: M.; Berger, C.; Cooke, R.; Grellert, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;
 submitted to the Protein Sequence Database, December 1999
 A:Reference number: Z23013
 A:Accession: T46059
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-547
 A:Cross-references: EMBL:AL132968
 A:Experimental source: cultivar Columbia; BAC clone T18N14
 C:Genetics:
 A:Map position: 3
 A:Introns: 24/3; 113/2; 188/1; 240/3; 350/1
 A:Note: T18N14.10

Query Match 6.5%; Score 702.5; DB 2; Length 547;
 Best Local Similarity 31.9%; Pred. No. 3.3e-14;
 Matches 189; Conservative 88; Mismatches 202; Indels 113; Gaps 17;

```

QY 195 EERSQOQDDIELETKAVGMSNDGRFLKFDIEIGSGFKTVYKGLDTETTVEAVMCELOD 254
DB 2 EISSASDDSIAYVEL-----DPSGRYGRFREVILGKAMKTVYKAPQVIGMVAAMNQVL 56
QY 255 RKLTKSER--QRFKEAEMLKGLQHPNIVRYFYDSWESIVKAKCIVLVTEIMTSGTLKTY 312
DB 57 NEVFRSPEDLRLYSVHLKTLKHSIKIFYSMIDHQMTIN-----LIEVFTSG 114
QY 313 LKRFKMKIYLRSGCQILKGLQFLHTRPTPIIHRDLKCDNIFITGPGSYKIGDLGLA 372
DB 115 RKRYKQVIRAKTSARQILNGLAYLHGHPPIIHRDLKCDNIFVNGHGLGVYKIGDLGLA 174
QY 373 TLKRAAS-PAKSVITGTPFMAPEMEYEEKYDSVDVYAFGMCLEMATSEVPSECCNAAQI 431

```

Db 175 AILRSGQNAHSYIGPEPMAPPELYEDVNLDIYSFGMCULEMLTGEYPSRCTNPACI 234
 432 YRRVTSVGPASFDFKVAPEVKEIIEGCIRONKDERYSIKDLNMAFQOETGVVELAE 491
 235 YKKVTSGLPDSFPHLIQHTEAQRFVGKCL-ETVSRRLPAKELLADPFL-----AATDE 286
 492 EDDGEKIAIKMLRIEDI-----KKLKGKKYKME-----520
 287 RDLAPLFRLLPQOAIQNLANGTVNHLBPSTDPRTTMDKSTIGKNSDEHTIFLOVOLI 346
 521 -----AIEFSPDLERDVEDVAQEMVESGVCEGDKHTAKAIXDRVS-LIKRRRQR 572
 347 DGGHNRNIQFPFNILSDTPELVALEMVKELEITMDPLEIAMINENISLIVPNRAND 406
 573 QLVREBQEKKKQEBSSLKQOVQSSASQTGIKQLPSASTGIPASTTSASVQVEEPP 632
 407 SSRHRE-----SFGHEDEDNDGTGRTLR-----FSSASSSHSDSPVAVREN 448
 633 EADQHQOLQYQOQPSISVLSGTVDSGQSSVFETSRVSSQQTIVSYG-----SQHQAHST 687
 449 NDS-----SNDVIPD--MDDGNRSS-----NRLNLSSTHYSPALDDQOQRRR 493
 688 GTVPGHIPSTVQAQSQ-----PHGVPPSSVAQSQSQSQSQSS 725
 494 VRLQQRKRSIVDRTRQVLRHSLMELINKRRGRGFDPTN-----ELQPOPSST 541

RESULT 5

TS1099
 MAP kinase [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cross)
 C/Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C/Accession: TS1099
 R/Cooke, R.M.; Berger, C.; Delzeny, M.
 submitted to the EMBL Data Library, May 1999
 A/Description: Analysis of Arabidopsis thaliana gene structure by cognate cDNA sequencing
 A/Reference number: 225297
 A/Accession: TS1099
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-549 <COO>
 A/Cross-references: EMBL:AJ238802; PIDD:CA43520.1

Query Match 6.5%; Score 702.5; DB 2; Length 549;
 Best Local Similarity 31.9%; Pred. No. 3.3e-14;
 Matches 189; Conservative 88; Mismatches 202; Indels 113; Gaps 17;
 195 EERSQOQDDIELEETKAVGMSNDGRFLKPIEIGRSFKTVYKGLDETTEVEVAMCELOD 254
 4 EISASDSDSLAVET-----DPSGRYGRFREVLDKGMKTVYKAFDVLGMEVAMNQVKL 58
 255 RKLTKSER--QRFKEAEMLKGLQHPNIVRPFYDSWESTVKGKCIIVLTELMTSGTLKTY 312
 59 NEVRSRPEPQLRTYSEVHLKLNLMHESIIRYCTSMIDV--NRRTFNFTLELFTSGTLREY 116
 313 LKRRKWKIKYLRKWCQIILKGLQPLHTRPTPIIHRDLKCNITPTGTSVKXIGDGLA 372
 117 RRRKQXVDIRAIKSWAQIINGLALYLGHDPPVIRHDLKCNITFVNGLGVKIGDGLA 176
 373 TLKRAS-FAKSVITGPEPMAPPELYEDVNLDIYSFGMCULEMLTGEYPSRCTNPACI 431
 177 AILRSGQNAHSYIGPEPMAPPELYEDVNLDIYSFGMCULEMLTGEYPSRCTNPACI 236
 432 YRRVTSVGPASFDFKVAPEVKEIIEGCIRONKDERYSIKDLNMAFQOETGVVELAE 491
 237 YKKVTSGLPDSFPHLIQHTEAQRFVGKCL-ETVSRRLPAKELLADPFL-----AATDE 288
 492 EDDGEKIAIKMLRIEDI-----KKLKGKKYKME-----520
 289 RDLAPLFRLLPQOAIQNLANGTVNHLBPSTDPRTTMDKSTIGKNSDEHTIFLOVOLI 348
 521 -----AIEFSPDLERDVEDVAQEMVESGVCEGDKHTAKAIXDRVS-LIKRRRQR 572

Db 349 DGGHNRNIQFPFNILSDTPELVALEMVKELEITMDPLEIAMINENISLIVPNRAND 408
 573 QLVREBQEKKKQEBSSLKQOVQSSASQTGIKQLPSASTGIPASTTSASVQVEEPP 632
 409 SSRHRE-----SFGHEDEDNDGTGRTLR-----FSSASSSHSDSPVAVREN 450
 633 EADQHQOLQYQOQPSISVLSGTVDSGQSSVFETSRVSSQQTIVSYG-----SQHQAHST 687
 451 NDS-----SNDVIPD--MDDGNRSS-----NRLNLSSTHYSPALDDQOQRRR 495
 688 GTVPGHIPSTVQAQSQ-----PHGVPPSSVAQSQSQSQSQSS 725
 496 VRLQQRKRSIVDRTRQVLRHSLMELINKRRGRGFDPTN-----ELQPOPSST 543

RESULT 6

T02951
 Probable mitogen activated protein kinase - rice
 C/Species: Oryza sativa (rice)
 C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
 C/Accession: T02951
 R/Kim, C.Y.; Cheon, S.Y.; Cho, M.J.
 submitted to the EMBL Data Library, July 1998
 A/Description: Identification and characterization of fungal elicitor responsive rice ger
 A/Reference number: Z14788
 A/Accession: T02951
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-677 <KIM>
 A/Cross-references: EMBL:AF080436; NID:G3450841; PID:G3450842
 A/Experimental source: cultivar Milyang 117
 A/Genes: MEK1

Query Match 6.3%; Score 678.5; DB 2; Length 677;
 Best Local Similarity 32.4%; Pred. No. 2.1e-13;
 Matches 176; Conservative 75; Mismatches 186; Indels 107; Gaps 14;
 218 GRFLKPDIEIGRSFKTVYKGLDETTEVEVAM--CELODKLTKSERQFKEAEMLKGL 275
 23 GRGRINDVLDGKASKTVRAPEYQMEVAMNQVKLHDFLQSPEDERLYCEIHLIKTL 82
 276 QHPNIVRPFYDSWESTVKGKCIIVLTELMTSGTLKTYLKRKYWKIKYLRKWCQIILKGL 335
 83 KHNIMKFTYSWDV--SRNINFTFMTFSGTLRQKRMVRNVIWAVKWCQIISGL 140
 336 QPLHTRPTPIIHRDLKCNITFPTGTSVKIIGDGLATLKRAKSAKSVIGPEPMAPPEY 395
 141 LYLSHDPPIIHRDLKCNITFVNGNQEYKIGDGLAAILRKSAAVHCVTPEPMAPPEY 200
 396 EEKYDESVDVYAFQMCULEMATSEYPYSECQNAQIYRRVTSVGPASFDFKVAPEVKEI 455
 201 EESTNELVDIYSFGMCULEMWTPEYPSCTHPVQIKYISGTRPALVKKVDPVNRQF 260
 456 IEGCIRONK-----DERYSIKDLNMAFQO-----ET 483
 261 VEKLTATASRLSAREVILKDPFLQVDLVFCPDGNGSNLMNYLRFQPLQHAYSTVSMNSN 320
 484 GURVEL-----AEDGEKIAIKMLRIED-----IKKLS-----514
 321 GLSIEDSDPTEDRWCEDDDIADGIDLFGHEDEBLGNDVITTIKRSSEDSITFLRL 380
 515 KYKQNEA-----IEFSPDLERDVEDVAQEMVESGVCEGDKHTAKAIXDRVS-LIKRRR 570
 381 RIANDGHVRNIYFPFDIADPTALSVALEMAELDIDHETVTRIAEMIDGVSLVDPWR 440
 571 QRLVREBQEKKKQEBSSLKQOVQSSASQTG--IKQLPSASTGIPASTTS--ASVSTQ 626
 441 PGGISQDPTTYGNCG-----SNVSSCGSLVAVMSCAARGCHCADLHGREEDITFQ 493
 627 VEPEPEADQHQOLQYQOQPSISVLSGTVDSGQSSVFETSRVSSQQTIVSYGSGHQAHST 686
 494 ANGEQDTLQD-----SGSSDDGGQOT--QHVADQEV-----HS 526

RESULT 7
T34434
Hypothetical protein K06A9.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C:Accession: T34434
R:Geisel, C.; Gattung, S.
Submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid K06A9.
A:Reference number: 221525
A:Accession: T34434
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2232 <GET>
A:Cross-references: EMBL:U08046; PIDN:AACT0890.1; GSPDB:GN00028; CESP:K06A9.1a
A:Experimental source: strain Bristol N2; clone K06A9
C:Genetics:
A:Gene: CESP:K06A9.1a
A:Map position: X
A:Insertions: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/1

Query Match 5.5%; Score 591.5; DB 2; Length 2232;
Best Local Similarity 22.8%; Pred. No. 2.5e-10;
Matches 385; Conservative 211; Mismatches 614; Indels 481; Gaps 72;

QY 587 SSIAQQVQSSASQTG-----IKQLPSASTGIPASTSASVSTQVEPEPEADQHQ 639
DB 593 SSSPQSSQSPABNTSTTPSQTSQSPSPSMN-PSSSTPTGSSQSTIPEG-----643
QY 640 LQVQQGISIVLSGCTYDSGGSSVFTESRYSQQTYSYSGHQHQAISTGVFGHISTVQ 699
DB 644 -----STASSPTSGTSGTFSVATE-VTSQSTVPSGS-----SLGT-----QSTN 681

QY 700 AQGQPHGVPPSS-VAQGGQSQGQSSSILTVSS--SOPLOHPQ-----741
DB 682 SSSPSSLSSTSGMTLISEPSSSTQSGAOSTITTPSPNSQSTSSLESSTGATTS 741
QY 742 --QGIQCTAPQ-QTVQVSLQSTSSSEATTQAQVSO--PQAPQVLPQVSAG--KQSTQ 794
DB 742 SCSAGTMTSPSSQSSVSGSSTSPASATTSCEMTSQSGTQTPGSSVSTSAAILNSTQ 801

QY 795 GVSQVAPAEVANAQQAQTPTLASSVDSAHSDVAGMSDGNENVPSSSGRHEGRTYK 854
DB 802 SVSTNSPGSTV-----TRPSTVSGSTSSG-STVVGSTEAATSGSSVASSSPAPSTQ 853

QY 855 HYKSVRS-----RSRHEKTSRPLRLINLVSNNKGRVVECOLETHNKKVTFKPLDD 908
DB 854 NRPSTSSSSSMTQTPYPSQSTSPVSSSTTPSGPGTTLTSTSSPQOSTITGTSQS 913

QY 909 NPEEIAITVNDFFILAIERESFVDQVREIIEKADMLSDVSVBEPGQGLSTLQK-- 966
DB 914 TSPGIST-----TSEMTSQG-STQPGSGTSTVTPSTV 947

QY 967 DDYGFSGQKLEGEFK--QTPASM-----PQ-----992
DB 948 SOSTSSGSTVTVGSTEAGSSPISTQNTNPSTSSGSSMSTQTPQSSQSTSPVESSTGA 1007

QY 993 --QIGIPTSLIQVHSAGRFIVSPVPSRLRESKVPSEITDVTAASTAGPG-----1045
DB 1008 TSSSGSPGTLTST-----SPSPSPSTSSSGSSTSPVSTISQSTETPGTST 1059

QY 1046 ---NLSHSAAGSLAQAFSELRRAQMTGEPYAPDNFSH-----TGPT 1086
DB 1060 VTKPSTVSGSASGATATMS--TEASTSGSSTSPNPSQSTSPSTGATSPSGSGTT 1117

QY 1087 FVVPPEFLSLAGVPTTAATAAPVATSSPNDI-----STSVQSGVTVPTLEGIAQVAT 1142

DB 1118 LTSISPSQSSSTIGSSQGSTSPVSTTS--GDMTSQGSTQIDPGSTSTVQCESTGSGST 1175
QY 1143 ST-GVATSGG---LPIPVSESPVLSVSVSSITIPAVVSTSTSPSLQVPTSTSEIVYSS 1198
DB 1176 STSGEITSGSTQTPPSSSLSTSPALSTSQ-----SVSTNSP-----GS 1215

QY 1199 TALPSEVTSANTSASAGSTATPGPPAVVSSQAAGSTVVG-ATLTVSVSTTSPSTAS 1257
DB 1216 TVTQPTVAGSTSS---GSTVITG-----STGSSSTSGSSSATSSSPVSTSQ 1263

QY 1258 QUSIQUS-STSTPTLAETV--VSAHSLDKTSHST--TGLAFSLADPS--SSPG 1308
DB 1264 SPNPSTSGSTTPPNSQSTSPVSTTGMTHSGSTQTPSTIGSTVTPQSTVSGNSG 1323

QY 1309 AGV-----SSYIQPGGLHPVLVPSVIASTPILPQAAGPT-----STPILP 1349
DB 1324 STVTTSSSASSTSGSSFKTSPSSISVPPTSSPPTTPASSTSGSTSDVSVSTTSLAP 1383

QY 1350 QVPSIPPLVQPVANVAVOQTLIHSQOPALLPNCPTHCEVSDSTQPKAPGIDIDIKTL 1409
DB 1384 LSSSLP-----STVPSSTQSPSTSESSKASSP--VPQSTSTPTNPSTSTSL 1434

QY 1410 EKLKSLFBEHSS-SGAQHAVSLETSLVIESTVTPGIFTTAVAPSKLLTSTTCLPT 1468
DB 1435 LSTTISGSTQHTTMSKASSGSTSPNSQSTGTVTWSGST--SGVSTSSASTQPM 1490

QY 1469 NLPLGTVALPVTPVTPGVSTPVSSTTSGVPRGTAPSKRPLKAVLPVGT--ELPAGT 1526
DB 1491 STSQSSA-----GSTVASSTASPAASSTAPSS-----TGWSSSTSGT 1529

QY 1527 LPSEQLPPEPPGSLTOSQOPLDLAQLRRLTSPXKIT-VTSAVGVSMAAPATIEAGT 1585
DB 1530 V-----GSLTSSS-----TTASASQSTSTVWSSSTSGVSTSSAST 1569

QY 1586 QPKGVSYQKEGVPVATSSGAGVFKKGRPOVVAADQAKE--GNKKEDAKSVHESS 1642
DB 1570 QPMSTSQSSASAGSTVASTAGVSTVPSSTGTWGSSTSGSTVSGSTTSSSTTASASQ 1629

QY 1643 T-----SESSVLSSSSESTVIVKEPENGITIPIS-----SDVPEE- 1678
DB 1630 TSGTVMGSSSTSGVSTSSAST--QPMSTSQSSASAGSTVASTTGLVSTVPSST 1685

QY 1679 -----AHKTJASEAKSDTGOPTKVRFO--YTTANKVRFSKTEBK 1720
DB 1686 GIMGSTSGTGSTTSESTASASQSTSTVWSSSTSGVSTSSASGQPMSTSGS 1745

QY 1721 ITDTKKEGVASPPFMDLEQAVLPVATPEKKEPELSEPHLNGPSSDPAAFLSRDVDG 1780
DB 1746 SAGSTVASTASP-----AASSTAP-----SSTGWTGSSSTGTVSGSTGASTA 1790

QY 1781 SGPSP--SPHOLSKSLPQNSQSLNSGFSNYSMSDNEIDEDDLKLELRRLDKHL 1838
DB 1791 STTSHSTGTVTLGSSSTSNQSTSQSSVSTVASS-----1827

QY 1839 KEIODLQSKHEISLYTKLGKVPNAVIIIPAAPLSGRRRPTKSGSKSSRSSSLGNK 1898
DB 1828 -----TAGLVSTST--VPSTGWTGSSST--VGSITSESTTAA 1865

QY 1899 SPQL-----SGNISGQSAAGVLAHQTLHPGNIPESSGQNOLOPLKPPSSDNLVYA 1951
DB 1866 SSQTSSTVMGSSSTSGVSTSSASTQ--PQWSTQ-----GSSAGSTVA 1908

QY 1952 FTSDGAIS--VPSISAPGQSTSTNTVGATVNSQAQOPAMPSTSRKGTPTDDLHKLV 2008
DB 1909 SSTAGIVSTVPS--STGWTGSSSTGUSTI--SESTIASSSTSQGS-----1954

QY 2009 DNWARDANLSGRRSGKHMNYEGPMARKFS-----APQQLCISMTSNLGSAP-- 2058
DB 1955 -----TVTIGSTSGT-----NPSSPRLSQITTPPQSTSTSTQSLSPSSSPST 2001

QY 2059 --ISAASATSL 2067
DB 2002 HSSVSSSEGTMM 2012

RESULT 8
T45025
muscin MUC5B, tracheobronchial (imported) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T45025
R:Gessy, J.L., Guyonnet-Duperrat, V.; Porchet, N.; Aubert, J.P.; Laune, A.
J. Biol. Chem. 272, 3168-3178, 1997
A:Title: Human muscin gene MUC5B, the 10.7 kb large central exon encodes various alternat
A:Reference number: Z22899; MUID:9716151; PMID:9013550
A:Accession: T45025
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-3570 <DES>
A:Cross-references: EMBL:Z22496; NID:91834502; PIDN:CA96577.1; PID:91834503
A:Experimental source: placenta
C:Genetics:
A:Gene: MUC5B

Query Match 4.6%; Score 494; DB 2; Length 3570;
Best Local Similarity 19.4%; Pred. No. 3e-07;
Matches 419; Conservative 226; Mismatches 705; Indels 814; Gaps 88;

QY 596 SSASQGIKQLPASATG-----IPTASTSASVTOVEPEPEADQHQOQOQOQPSISV 650
DB 1177 TTTTATSVTPPISSSLGTTTTLRLSQTTPATMTATPSTPETAH-----TSTVL 1227
QY 651 SDGTVDSGQSSSVFTES-----RVSSQOTVSYGQHEQAH-----STGTV 691
DB 1228 TATATTATGATGATPSTGTATHTKVPPTTTTGTATPSSPGATLPFWISTTTP 1287
QY 692 GHTPTVVOQASQPHGV-----PPSSVACG-----GQOQOPS--SLTGVSS 733
DB 1288 TTRGSTVTPESIGTHTATVLTITTTTATGSMATPSSSTQTSQTPPSLTITATITAT 1347
QY 734 OPTIHOQOQGIQOTAPPOQTVQYSLQSTSEATTAQVSOQAQOVLPOVSAGKST 793
DB 1348 GSTTNSSTGTGTPPIPVLT--ATTPTATSSVTPSSALGTHHPV--PNTTA--TT 1401
QY 794 QGVQVAPAEVAV-----AQPOATPPTLASSVDSA-----HSDVASG-- 832
DB 1402 HGRS--LSPSSPHVTRATMTATSGLTGLTHTEPSTGSHTPATGTTGHTSPALSPH 1460
QY 833 -MDDGHNVSSSGRHGR--TKRHKRSVRSR-----HEKSRPLRLINS-- 880
DB 1461 PSSRTTESPPSPGTTTPTGHTATSRRTATATPSKTRTSTLLPSSPTSAPIITTVMGCEP 1520
QY 881 -----NKGD-----RV 886
DB 1521 QCAMSEMLDYSYMPGSGSDPDTYSNIRAGAVCEQPLGLECRAGAOQVPLRELQV 1580
QY 887 VECQLE---THNRKV---TFKEDLD-----GDNPEIATIMVNDPILAIERESF 931
DB 1581 VECSLDFGLVCRNREQVGFKMCENYERIVCCNGVCFSTPAT----- 1624
QY 932 VDOVREIEXADEMLSDVAVPEBGDQ-----LESIQGCDYGFSSQKLEGEFKQPI 986
DB 1625 -----SSTATPSPGTWILTEQTTATTTATGTTATPSTPBTAP 1667
QY 987 ASS-MEQOIGIPTSLSLOVHSAGRPIV-----SPVESRL----- 1022
DB 1668 PRVLVSQATTPATSKATSSSPKATTLPLVLTSTATSTATSPFIPSSILGTTGTS 1727
QY 1023 -----RESKVPSEITDVAASTAGSPGMNLSHASSLSLQQA-----FSELR 1066
DB 1728 ONRPPHMTATSTIHPESTPETHSTVLTAKATTATATSMSTPSTPCTWILTELT 1787
QY 1067 AQMTBG--PNTAPNHSHTGPTFPVPPFLSSIAQVTTAAATAPVATSPENDISTSV 1124
DB 1788 AATTALPHGTPS--STPGTWILHEPSTTATVTPGTSTATAS----- 1830

QY 1125 IOSEVTPTEEGIAGVATNGVTSGLPIPPVSESVLSSVSSITIPAVSIS----- 1179
DB 1831 -STRATAGTLKVLSTATTATPVISRAPSSSGATLALPALMSATATTPATSVTAIPSS 1889
QY 1180 -----TTSF-----SLQVPTSTSEIVASTALYPSVTVASASAGSTATPGK 1224
DB 1890 SLGTATRLSQTTPATMTATPSTSPTEVHTSTVLTATTTATTTT-----GSVAPSS 1944
QY 1225 PRAVVSQQAAGSTTVATLTVSTTTSFSTASQSLQSSSISTPLAETVVSASHL 1284
DB 1945 PGTAHTTKVPTTTTFT--ATPSSPGTALTPVWISTITPTTSGSTVPS--SIP 1998
QY 1285 KTSN-----SSTGLAFSLAPSSSSPGAGVSSVISOQGLHPLVPSVASPILP 1337
DB 1999 GTHTATVLTITTTTATGSMATPSS-----TQTSQ-----TPPSLTATTTI 2042
QY 1338 QAAGPTSTP-LPQVPSIPPLVQVAVNPVAVQTLI-----HSQOPQ----- 1378
DB 2043 TATGTTNPSSTPGTPIPPVLTATTPATATSVTPSSALGTTTTPPVNTTATTHGR 2102
QY 1379 ALLNPQHT-----HCPEDVS--DTPPKAGIDIDITLEE 1411
DB 2103 SLPPSPHVTPTAMTATSGLIGTHTTEPSTGSHTPATGTTQPSPT----- 2152
QY 1412 KLSLRFSEHSSGQAQASVSLFSLVIRESTVPG-----IPTAVA--PSKILSTSTCL 1465
DB 2153 -ALSSPHSSKTTSPS-----PGTTTGHGRGTSRTATATPKR--TRISTILL 2199
QY 1466 P--PYNLPLGTV----- 1475
DB 2200 PSSPTAPITTVVTTGCEQCAMSEMLDYSYMPGSGSDPDTYSNIRAGAVCEQPLG 2259
QY 1476 -ALPVTVPVTPGOV-----ST 1490
DB 2260 LECRAQAPGVPLRELQVVEGSLDFGLVCRNREQVKEMCENYERIVCCNGHCPST 2319
QY 1491 PV--STTSGVKRG--APSKPLTKAPVLPV-----TELPAG 1525
DB 2320 PATSSATPSSPTGTWILTKLTATTTSTSTSTPSTQGP--PAGPHVSTIATTP 2377
QY 1526 TLPSQLPPPGPSLTQSQOPLDLDLQRLTSPKXITVTSAGVSM-----A 1575
DB 2378 TVTSSKATPSSPG--TATALP-----ALRSTATPTATSTFALPSSSLGTTWRLSQT 2430
QY 1576 APTAIIEAGTQPKGVSQVKEGVPVLTSSGAGVFKMGREQVSAAGAOKGKNKSEDAK 1635
DB 2431 TPAATMSTAT--PSTPEVHTSTVLTITTA-----TTGAGAGVAATPSTGTAT 2478
QY 1636 SVHFESTSES--SVLSSSPE-----STLVKEPNGITIPGISDVPESAHKTTAS 1685
DB 2479 TTKVPTTTTGTPTVSSSFGTARTRPFWISTTTTPTGSGTV--TPSSIPGTHPT-- 2534
QY 1686 EAKSDTGOPTKGRF-----QV 1702
DB 2535 -VLTITTOVAGSMATPSSSTQTSQTPPSLTITATTTATGTTNPSSTPGTPIPEL 2593
QY 1703 TTTANKVGRFVSXTEDKLTIDTKEGVPVAPPMW-----LEQVILPA 1745
DB 2594 TTTAT-----TPATSTVTPSSALGTHHPVPPNTTATTHGSLSPSHVTRATMSA 2648
QY 1746 VIPKKEKPELSEPSHNLNGPSSDEEAAFLSRDVDGSGSPHSPHQLS----- 1791
DB 2649 TSGTLGTHHTPS--TGTSHTPAAT--TGTTTSPALSSPHSSRTTESPPSPGTTTP 2704
QY 1792 -----SKLPSQNLSQL-----SNSFNSY--MSD 1816
DB 2705 GHTTATSRRTATAPSKIRNTLLPQOPSAPITTVTTGCEBQCAMSEMLDYSYMPG 2764
QY 1817 NESDIE-----DEDKLELR-----LADKLEKEI 1841
DB 2765 SGSDPDTYSNIRAGAVCEQPLGLECRATAQGVPLGELQVVEGSLDFGLVCRNREQ 2824
QY 1842 QDLQSRQKHEIBSLYTKLQVP--PAV----- 1866

Db 2825 GKFMKCRVYELRVCNCGHCPSTPATSSSTPSTPCTTWTLELTATTTATTAAGTSTA 2884
 QY 1867 -----IIPPAALSGRRRRRTKSKGSKSSRSS--LGKSPOLSGNLSCQSAAYLH- 1916
 Db 2885 TPSSTPGAPRPVKLTSPATPT-ATSSKATSSSSPKRTATTLPLVLTSTATSTSTVTPPI 2943
 QY 1917 PQQLLHPHNGINPESGQNL-LQPLKPSSSDNLYSAFTSDGALISVPSLSAPGQSTSTN 1974
 Db 2944 PSSLTGTTGTLPEQTTTPVATMTSIHPSSTPETHTS-----TVLTATKATTAATSTTS 2996
 QY 1975 TVGA-----TVNSQAQAQAPAMTSSSKGTFPTDDLHLVUNMAPADAMNLSC-R 2021
 Db 2997 TPSSTPGTTWLTLELTATTTAGTGPATPSTPTGT-----TWLTLELTATT 3046
 QY 2022 RGSKGNHNYEGPMARKFASAPQLCI-----SMTSNLGSAPISASATS-LGHFTK 2072
 Db 3047 TATG-----STATLSTPFTWTWLTETSTRTVTAPEGSTATTAATAGTTPHVT 3099
 QY 2073 SKCPPOQYGPATPFGAOWSGTGAPAPQLGQFQPVGTASLQNFENISLQKSIINPGSN 2132
 Db 3100 TATTPVTSSKATPSSSPTATLPA-----LRATTPPTATSPATPSSS 3145
 QY 2133 LRTT 2136
 Db 3146 LGTT 3149

RESULT 9

S69703
 HKR1 protein precursor - yeast (Saccharomyces cerevisiae)
 N/Alternate names: protein YDR420w
 C/Species: Saccharomyces cerevisiae
 C/Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
 C/Accession: S69703; A53382
 R/Dietrich, F.S.
 submitted to the EMBL Data Library, August 1995
 A/Description: The sequence of S. cerevisiae lambda 3641 and cosmids 9461, 9831, and 941
 A/Reference number: S69555
 A/Accession: S69703
 A/Molecule type: DNA
 A/Residues: 1-1802 <DIE>
 A/Cross-references: EMBL:U03007; NID:g927685; PIDN:AMB6457.1; PID:g927691; MIPS:YDR420w
 R/Kasahara, S.; Yamada, H.; Mto, T.; Shiratori, Y.; Miyamoto, C.; Yabe, T.; Nakajima, T.
 J. Bacteriol. 176, 1488-1499, 1994
 A/Title: Cloning of the Saccharomyces cerevisiae gene whose overexpression overcomes the
 A/Reference number: A53382; NCID:94156857; PMID:8113191
 A/Accession: A53382
 A/Molecule type: DNA
 A/Residues: 1-581, 'A', 583-593, 'A', 595-1802 <KAS>
 A/Cross-references: EMBL:S69101; NID:g545659; PIDN:AMB30051.1; PID:g545660
 A/Experimental source: YNN295
 A/Note: sequence extracted from NCBI backbone (NCBIN:144410, NCBIP:144411)
 C/Genetics:
 A/Genes: SGD:HKR1
 A/Cross-references: SGD:S0002828; MIPS:YDR420w
 A/Map position: 4R
 C/Keywords: calcium binding; glycoprotein; transmembrane protein
 F/1-21/Domain: signal sequence #status predicted <Sig>
 F/22-1802/Product: HKR1 protein #status predicted <MAT>
 F/1483-1508/Domain: transmembrane #status predicted <TM>
 F/1645-1656/Domain: calcium binding #status predicted <CAL>

Query Match 4.4%; Score 478; DB 2; Length 1802;
 Best Local Similarity 2.6%; Pred. No. 4.3e-07;

Matches 333; Conservative 209; Mismatches 604; Indels 328; Gaps 64;

QY 596 SSASQGIKQLPQBSAGTPTASTSASVTOVEPEPADQHQOL--QYQPSISVLSLSD- 652
 Db 29 STSYNNGISTESTSYST-----SAISSTGSNKENATTSSESTTTWAGQVGGESSTIMDE 83
 QY 653 -GTIVSGGSSVFTTERVSSQCVVSGSHQAHSGTVPGHIPSTVQAQSGPHGV---- 707

Db 84 QETGTSQVYISVTTTQTSPTDMS-SVKSTELIATPSSSI---VEPLQSVSEDSQISQTL 139
 QY 708 -YPPSSVAQSQSQQP-SSSLTGVSSQCIQHQQOQGIQQTAPPOQTQVYSLSQSTS 765
 Db 140 SHNPKSAEEDSDTSTSESSSVIITSDDSAVREISPTITTDQSISKEGTLAQTSI 199
 QY 766 SEATTAPQVQAPQVLPQVSAAGKQSTQVGVAPAEVAVQAPQAPQATPLASV--- 822
 Db 200 SEIT-----KLAQVMTVSG-----ISSTIASTIDGSESTQ-TDFSNVSPFE 243
 QY 823 DSAHSDVASGMDGNNENVPSSSGRHGRTTKRHYKRSVRSRSHREKTSRPLRLIANYNK 882
 Db 244 NSVEEYAMSKSQLSSEYSSSSVYSG-----GESTADKTSSSP-ITSPSSS 289
 QY 883 GDVVEQQLETHRKMTFEPDLGDNPEDALIMNNDLILAEKRSFPDQVEITEKA 942
 Db 290 YSQTTSTETSESSRVAV-----GVSRPSSITQTTSI----- 320
 QY 943 DEMISEDVSVPEGDQGLSELGKDPYGFSGQKLEGEFKQPIPASMPQOI-----GI 996
 Db 321 DSRMSREVELSTY-----YDSAGVPPDELIVDRPATSTETSEASQGV 367
 QY 997 PTSSLQVYHS-AGRPVSPVESRLRESKVPSEITDVAASTAQSPGMNLSHSASL 1055
 Db 368 SRESNTEFAVSSISTTFEIVSASDT-----VYSTSTNTVPSVSHSTVHAATSSSTYI 421
 QY 1056 SLQQAFTSLRAQMTGPNTPAPNPFSGTGFVVPPLS---SIAGVPTTAATAPVPA 1112
 Db 422 S-----SSLYSPSLSASVSHFG-----VAFPPAYISFSSVPVAVSST---Y 462
 QY 1113 TSSPPNDISTVYQSEVYPTTEGIAGVATSTGVNMSGPLIPVPSPPSLSSVSIIT 1172
 Db 463 TSSP-----SAQVVP-----SAYAS-----PSVPA---VSTYTSPPSA 496
 QY 1173 PAVVSTIS-TTSPSLQVPTSTSELIVSS-----TALYPSVTSATSASAGSTATPGKPPA 1227
 Db 497 PAIISSTYSSPS--APVAVSSTYSSPSAPAIISSTYSSPSAPVAVSSTYSSPSAPA 554
 QY 1228 VVSGQAAGSTYVATILTSVSTTSPSTASQLSIQSSSTSTPTTAEVTVVVAHSLDKTS 1287
 Db 555 AISTYTSPPS--APVAVSSTYSSPSAPVAVSSTYSSPSVAVAVSSTYSSPSAPAI 612
 QY 1288 HSGTTG-----LAFSLAPSSSSSPGAGVSYISQPGJLHPLVTPVIASTPTLPQAPG 1342
 Db 613 SSTYTSPPAPVAVSSTYSSPSAPAIISSTYSSS--VPVAVSSTYSSPSAPAIIS 670
 QY 1343 TSTPLLPQVPSIPPLQVY-----ANVPAVQQLIHSGQOPALLPQPHTHCEVDS---- 1394
 Db 671 TYT-----SSPSVPAVAVSSTYSSPSAPAIISSTYSSPSAPVAVSSTYSSPSAPA 726
 QY 1395 ---DTQPKAGIDDKTLBEKLSLPSEHSSGQAQHASVGLSVLESTVYTGIPPTTAVA 1452
 Db 727 TYTSSPSAP-----VAVSTYSSPSAPAIIS-----STYTS--PSAPA 765
 QY 1453 PEKLLTSTSTCL-----PTNLPLGTVALPVT-PVVTQGVSPVSTTSGVKPGAPAK 1507
 Db 766 VSTYTSSSALVLSSTSTSSPYDIYSPSTFAAISSTYTS--SASVAMSTSSSS 822
 QY 1508 PLVTAAPVLPVGTLPAGTLPEQLPPPEPSLTOGQPLDLDAQRLRLTPEKITYVS 1567
 Db 823 P-----YDIYVLSASSARSISLATEFSPSPSTLPTS-STYTFSSAYAREFSERSVTS 878
 QY 1568 AVGPVSM-AAPTAITEAGTQPKGVSGVKEGVLAT-----SSGAGVFMKGRQVSV 1618
 Db 879 TIAPIQHSTLRIIDPFLQTSMAIQSVISQISTSTLNDIHSALSVF----- 929
 QY 1619 AADGAQKEGKNKSEDAKVHFESSSTSESVLSSSEPTLVPEPAGTIPGISDPVPS 1678
 Db 930 -----NPS-----ASNLTETSLIIST-QAETIPK-NSAKISSLQQLSSS 969
 QY 1679 AHK--TTASEAKSDTGQPTKVGKRFQVTTTANKVGRFSVSKTEDKITDTKKEGPVVASPPPM 1736
 Db 970 TKNPYDTANKTETGRGTGVVSNFLYTSAAKPDNEKESATPTETI----- 1015

QY 1737 DLEQAVLPVAVIKKKEPELSEPSH---LNGPSSDPEAPLSPDVDDGSGP---HSPHQ 1789
 DB 1016 -----TISSSSHAYSLSPSSSHVNTGLSHNFVSSKATSFYSSSS 1058
 QY 1790 LSSKSLPSQNL--SOSLSN-----SPNSYSSNDSESDIEDDLKLELRLDLKEL 1841
 DB 1059 ISSIKLSKETTPASKSVSNQERITSTLTLAN---SQSEKSEGNNSVGLSSSHSSN 1115
 QY 1842 QDLQSKHEIESLYTLGKVPVAVIIPPAAPLGGRRRPTKSKSKSRSSSLGNKSPQ 1901
 DB 1116 PSLSTWTKVDSKSLSRKSVSKTNGE-----NGEETGLTTKTQYKSSSETSGSYSRS 1166
 QY 1902 LSGNLGGQASAVLHPQGLTHPPGNTIPESGONLOLPKSPSSDNLYSAFTSDGALSP 1961
 DB 1167 FT-KISIGPTTAVQGTASTNSVFTAPALS---TYFTTYP-SPNSYAMLPT--AlIVE 1218
 QY 1962 SLGAPGGSSTNTVGTATVNSQAAPAMTSS 1995
 DB 1219 S-SETGETTASFN---PSITGSLPMLEPVAVS 1248
 RESULT 10
 S26058
 Probable transforming protein (can) - human
 C/Species: Homo sapiens (man)
 C/Date: 13-Jan-1995 #sequence-revision 13-Jan-1995 #text-change 08-Oct-1999
 R/Accession: S26058
 R/von Lindern, M.; Fornerod, M.; van Baal, S.; Jaegle, M.; de Wit, T.; Buijs, A.; Grosva
 Mol. Cell. Biol. 12, 1687-1697, 1992
 A>Title: The translocation (6;9), associated with a specific subtype of acute myeloid le
 k-can mRNA.
 A/Reference number: S26058; MUID:92195315; PMID:1549122
 A/Accession: S26058
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-2090 <YON>
 A/Cross-references: EMBL:X64228; NID:g29652; PIDN:CAA5535.1; PID:g29653
 Query Match 4.4%; Score 475.5; DB 2; Length 2090;
 Best local similarity 20.7%; Pred. No. 5.9e-07;
 Matches 392; Conservative 238; Mismatches 740; Indels 525; Gaps 79;
 QY 545 YVCEGHKTKAKIKORVSLIKRREO-----RLVREHGGKKKQEGSSSLKQVEGS 596
 DB 309 YIEWDVLVAASASTEVSLARQSDQINNESWLEDSRAELPYTKSDSLPMGVVD 368
 QY 597 SASQTGI-----KQLPSA-----STGIPTASTTSASVSTQVE--- 628
 DB 369 YTNQVELTIDKTLPPAPVIMLSTDGYLCFPMINQNGVMSLITPELSLEGROP 428
 QY 629 -----DEEPPADHQO-----LQYQPSISVUSDGTVDSGQSSVFT---ESR 668
 DB 429 KSPGSTPTTPTSSQAPKLDASAAAAPASLPSSPAPPIAFISLLPAGAPVTFSSSS 488
 QY 669 VSSQGV-----SYGSGHEQASHGTVPGHIPSTVQAGSQPHGVPPSSVA---QGSQ 719
 DB 489 LKSAATVTGEPSTSSGSD--SKAAPGPEPST-----FSFVPPSKSLATTPASP 538
 QY 720 GQSSSSSLTVGS-----SQPIQHPQOQO-GIQQTAPPO-QTVQVSLQSTISSAAT 770
 DB 539 VAPPAASFSSGSGFKPTLESTPVPVSAPNIAKSSPSTSAVKVNLSEKPTA--AAT 596
 QY 771 AQVPSQPO-AQVLPQVSAGKOSTOG-VSQVAP--AEVVAQAQOAPQOPTLASSVDSAH 826
 DB 597 STPVSSSQSAPPMSPFSSASKPAASGSLPPLSAPPSSVPLKSVLPSPSGSAQSS 656
 QY 827 SDVASGMSDGENPSSSGRHEGRTTKHYKKSVRSGRHE-KTSRPLRLTLNNSKGR 885
 DB 657 SPVSSMQKSPRTTPPA--KPSPPQAKSLQPAVAEKQGHQWKSDPVM-----AGIGEE 709
 QY 886 VVECOLTHNRKMTTFKFDLDGNDPEELATIMVND-----FILAIER----- 928

DB 710 IAHFQKEBELKARTSKACFOVGTSEEMKMLRTESDGLHFILEIKETTESLHGDISSLK 769
 QY 929 ---ESF---VDQVREILIEKADEN-----LSEDSVVEPEGQGLSLOGKDDYGHSSGOK-- 976
 DB 770 TLLIEGFAGEVEERENENRDSGYLHLYKREPLDKSEQLOEIRLHOYKFAVQDVN 829
 QY 977 -LEGFKQIPIPASSMPQOIGIP-----TSLTQVWASAGRRF--IYSPVESRLRE 1024
 DB 830 DVIDLEMDQHEQKKQRHLVBERETLNTLANNELIINQCKRLNHLVDSIQQLRLYK 889
 QY 1025 -----SKVPSSE-----ITDTVAASTQSPGMNLSHSAASLSLQQA9SELR 1065
 DB 890 QTSIMELSSAVPEQSSIHFSFDSLESLCNALKTIE-----SHTKSLPKVPAKLSPMK 943
 QY 1066 RQAM-----TECP---NAPPNESHGTGPRFPVPPFLS-----SIAGVPTTAAT-- 1107
 DB 944 QAOQLRNLAKRKTPPVRSAPASLSRSA-----FLSGRYEDDESVSTSSVSGLE 995
 QY 1108 -----APVATSSPNDISTSVI-----QSEVTPTEEGIAG- 1139
 DB 996 SEDARTSCDDEAVQAPRHAPVVRTPS---IQSLPLHAPAPFAKSHLVHSSPGVMGT 1051
 QY 1140 -VATSTG-VVTSQ-----GLPIP-PV-----SESPVLSV 1166
 DB 1052 SVATSSAKIIPQADSTMLATKTVKHGAPSPSHISAPQOLAAALRQWASQAPAVNTL 1111
 QY 1167 VSSI--TIPAVVISITTSPELQVP--TSTSEIYVSTALYS---VTVASATASAGSTAT 1220
 DB 1112 TESTLKNVPVAVVQELKNNPATPSTMGSSVSYSTKTHPVLTPVANAQAGSLINS 1171
 QY 1221 PGKPPAVVNSQAAGSTVAGATLTSVSTTSPFSTASQSLQSLSSSTPTLAEVTVVA 1280
 DB 1172 LKESGTPPAGQSLSDPKASGAKIEAVITSTBSAQSFKPSFSGFNGFIITPT 1231
 QY 1281 HSLDKTSHSSTGLANSLAPSSSSSPGAGVSYISQPGELHVLVPSVASTPIIPQA 1340
 DB 1232 PSSNFTAAQAGTSTKSSQPDAPSSGSGSKPEYEAIPSSPSGGLTASANTTPGEPAAS 1291
 QY 1341 GPTSTPLPQVPSIPLVQPVANVPA--VQOTLIHSQOPALLPNQPHTHCEVDSPTOP- 1398
 DB 1292 --SSRPVAPSGTSLSTSSKLETPPSKLGELPPSSLAGTSLSGLRGQADDSKPT 1349
 QY 1399 -KAPGIDIDKLEBEKRLSPSEHSSG---AQHASVLETSIVESTVTPGIPPTAVAP 1453
 DB 1350 NKASSTSLSTQPTKNSGVSGFNFTAPVLGKHT-----BEPVTSATTSVAP 1399
 QY 1454 SKLITSTTCLPPTNLPGLTVALPVTPVTPQVQV---TPVST-----TISGVK 1500
 DB 1400 -PATSTSTSTAV-----FG--SLPVTSAGSSGVISFGTSLAGKTSFSGSQOINTSV 1450
 QY 1501 PGTAPEKPELTKA-----PVLPGTELPAGTLPSBQLPPEFGPS---LTQSQQPLEDL 1550
 DB 1451 PPSAP--PPTTAATPLPFTSPFTSLFSGLSLSTTPS--LMSAGRSITBEATSLPPEKPG 1506
 QY 1551 DAQLRRLTSEKXITVTSAVGVMAAFTAITAGT-----QPOKGVQVKEGVLIATSSG 1605
 DB 1507 DSE-----VSASASLIEEOQSAQLPQAPQOTSVDYKKEPVL----- 1543
 QY 1606 AGVFYKGRFOVSVADAQAGEKKNKSDASVHPESTSSSVSSSPSS--TLVKEPN 1664
 DB 1544 -----AQAVNSGTAASSTSLVALSAEATPTTVPDPAFEAVPPAS 1566
 QY 1665 GITIPG-----ISSDVESAHTKTAASEAKSDT-----GQPKVGRFOVQTTANKV-- 1709
 DB 1587 SFSVPGQTAVTAAAISSAGPAVAETSTPTLASSITSLVAQCPBSAALFGIVISGSSVFA 1646
 QY 1710 --GRFSVSKTEDKITDTKKGFPVASPEPMDLEQAVLPAVTPK-----EXPELSEP 1758
 DB 1647 QPAAASSSAFNGQNTNNTATAPSATPVFGVAASTASLSLFOQGTGSTATAATAPVSS- 1705
 QY 1759 SHLNGPSSDPEAPALSLSDVDDGSSPHSPOLSKSLPSQNLQSLNSNFSSTSSGDN 1818
 DB 1706 SGFSSPAFGTTAPGVQGTTEGQASV-----FGQSASMAASVSPFSQPGF 1750

QY	1819	SDIEDDELKLELRNRKRIKHEI	QDLOSROKHEIESLYTKGVP	PPAVIIPPAAPLSGR	1878
Db	1751	SSV-----		PAFGPPASS-----	1762
QY	1879	RREPTKSGS-----	KSSRSSSLGNKSPQLSGNS	SGSASVLPHQOLHPPGNI	1930
Db	1763	-TPTSTSGSVFGAASSTSSSS	SFSFGOSSPNTGGGLFGOS		1806
QY	1931	GQNQLIQLPKPSPSSDNLY	SAFTSDAISVPSLS-----	APGQTSSTNTV-----	1976
Db	1807	GQ-----	SPFGQGGSGVFGGTS	SAATTATTSQSFSGFSS	1857
QY	1977	-GATVNSQAQAQAPP	PMASTSRKCTFTDHLKLV	MDNARDAMNISGRGSK	2035
Db	1858	GGIVTFQGGSSSSSGSV	FGSGNTGRGGGFGFGSG	CGKAPCDANKNPSSAS	1910
QY	2036	ARKFSAPGQICISMT	SNLGGSAPIASAATSL	GHPFKSNCPQGYFPAT	2095
Db	1911	-----	GSTATSNSTNLFGN-----	SGAKTGGFASSSFG	1952
QY	2096	GAPAPLGGPQPVGT	ASTLQNFNISLQKSI	SNPPG 2130	
Db	1953	SVASQGFSSSPKKTG	FGGAPVFGSPPTFG	SGPG 1987	
RESULT 11					
T48814					
hypothetical protein 15E6.220 [imported] - Neurospora crassa					
C:Species: Neurospora crassa					
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #ext_change 05-May-2000					
C:Accession: T48814					
R:Submitter: U; Aigm, V; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura					
submitted to the Protein Sequence Database, April 2000					
A:Reference number: Z24541					
A:Accession: T48814					
A:Status: preliminary					
A:Residues: 1-1952 <SCH>					
A:Molecule type: DNA					
A:Cross-references: EMBL:AL353822; GSPPB:GN00112; NCSP:15E6.220					
C:Genetics:					
A:Gene: NCSP:15E6.220					
A:Map position: 2					
A:Introns: 281/3					
Query Match					
Best local Similarity 4.4%; Score 474.5; DB 2; Length 1952;					
Matches 446; Conservativeness 238; Mismatches 748; Indels 699; Gaps 101;					
QY	6	AEKQSTSGSLFLSP---	PAPAPKMGSSSD--	SSVGEKLGAAADAVIGRT	60
Db	166	AADQPKINGIAHSNPTNR	PNTDLDKGRDTPQVTSV	---ATGATVG-----	209
QY	61	MDKDSGAATTTTTEH	RFRRSVICDSNATAL	ELPGLSLPQPSIPAAV	120
Db	210	-----GAAATTVP	GN-----	STNSTSHFGGLMDGP	247
QY	121	REETVTAATTSOVA	QAPPAAAPGEQAV	AGAPSTVPSSTSKDP	171
Db	248	-----	AVDGMANNDFPST	MDLLOSOGHILMSF	280
QY	172	VGSKEPPPARSGG	SGAKPEQERSQOQ	ODIELELEKAVGMS	231
Db	281	SGG-IQPPPAATA	ADASIESPARLE	FADGVQMTTIALI	327
QY	232	FKTVYKGLDIT	VEVAMCELODRK	LTKSER-----	281
Db	328	-----	AW-----	FLAKKEERRAQY	365
QY	282	RFYDSW-----			287
Db	366	RFSSSYISIEB	GMGPSSDDEDE	BERGPPSDNRAP	425

Qy	288	ESTVKGKCIIVATLTM--TSGTLKYLYKRF-----	WAKIKYLRBW	327
Db	426	DQTVDDGKGPNRQYVSHTPGAAVNLALRPSPHHTFELGHSBQPNIAATKA1---	482	
Qy	328	CROILKGLQFLHT----RTPTIHRDLKCDNFIPTGTSVKIGDGLATIKASF----	379	
Db	483	SRHHLK-IDPNSQAGFEA1PLHKNGFFEDHNYHSDKYVLKSGD--RLQVKVEVEFII	539	
Qy	380	---AKSVIGTPEPMABEMTEEKYDESVDVYAFGCMLEMATSEYPSSECONAAQIYRYVT	436	
Db	540	NGVAEBKTAEEYEPETPARRYSEGGKEMSF-----	DESIHDVDRSGI	584
Qy	437	S---GVKPAFPOKVA1PEVKEI1EGCIRONKDERYSIKOLLNHAFFOESTGYRVELAED	493	
Db	585	SPEDNENVPMSVD---PD-RDQSQELSPREDIILLPAIPHVM--BT-VEKEAEED	635	
Qy	494	DGKTAIKLMLRIEDIK-KLKGRKYKDNEA1EFSFDLERDVPEDVAQEMVESGYVGCDK	552	
Db	636	DDEEAR---SOSQSVKPRPEPFPMDSVPL-----MDPRKX-----	RGPBR	674
Qy	553	TMAKAIKDBVSLIKRRBEQOLVREBEKKKEBESLKQV-----EGSSASQIGI	603	
Db	675	PEKNGIMSGREBRRLKQOMELAKKNQRPORGPGEPRVKKVRPRKHPLFEDAPDRPEKR	734	
Qy	604	KOLPSASTG1PTASTTSASVSTOVEPEEP-----EADHQOQLQYQPSISVLSDT	654	
Db	735	KYKPRKQNEEGDASDAEKTIKERKREKTPPLERLREDYTBEOQLKPKRYGUILDEV	794	
Qy	655	VDSGGS-----SVFTESRVSSQQTAVSYGS---QHEQASHGTVP	691	
Db	795	LSAABDGLTKQIYKRIOLKYPFYFYNVDYTKMESSVHNHILGNDAFKNEETHLSRVP	854	
Qy	692	G-----H1PSTVOAOS-----OP-----HGV---YPRSSVAQGS	718	
Db	855	GIDIDAGKKRKARPSDHA5SLHNFGQHYAPDRMPRPHMGHGHGBVQOSYHNGTGVQORS	914	
Qy	719	---QGQSSSSILTVSSSQPIQH-----PQG-----OQGIQOTAPPOQTVOY--	757	
Db	915	YVTTQQPG-----ASQHPQHILQTPQGVVPQGRPRAYQAAQTSPPRAQOYUGRPE	967	
Qy	758	SLSTGTSSEATQCP--VSOQAPQOVULPOVSA-----	GKOSTGVSYQAP	801
Db	968	TAARQMSGTPRAVYTSPPVSRPMPETVAAGAGATPHSMARQHSILPVSQPSQANGIPIRVNP	1022	
Qy	802	AEPVAVAPQAOPTPTLLASVDSASHDVA5QMSQGENEVPSSSGRHEGRITK-----	853	
Db	1028	-----PRATANTGV-----PAGAGARPAQOATPRANTTAPRNLNPIVLAPELIS	1070	
Qy	854	--RHYKSVBRBSRHEKTSR-PKIRILNVSKNGRNVVEQULETNRKXVTFKFDLDGDP	910	
Db	1071	WLESKATVEKLEYIQKTSKFPQILAMSVNIRG-----LKLTKSMI-----P	1113	
Qy	911	EEIATIMWNDFILAIERESFVDQVREIIEKADEMLSEDEVVEPEGGQGL-----ES	962	
Db	1114	DEESILEY---VLRV---FEERIQGSHK-----SLDDLIQTLILTFKATWST	1156	
Qy	963	LQKDDVCFSSQKLEGEFKOP1PA5SNPQOIG1PTSSGLTVVH5AGRRFTVP--VPE	1019	
Db	1157	LEAKID-----SQRAE---CIVLSAIDQVLGLADKTI1FRGTSESEMEFEFNNAKVLIPA	1206	
Qy	1020	SLRESKVFPEIRIDTVAA--STAQSPGMLSH-----SASSLSLQOAF-----SELRAPQ	1069	
Db	1207	IRMYAEHQKQVAA1TAAPRPHATITGALPANHTHTAP1APSAAQRIATPSATORANP	1266	
Qy	1070	TEG--PNTAPNFSHTGTFPVVPFL--S5IAGVP-----TTAATA-PVPA-----	1112	
Db	1267	PGGNPSAVPRTITTAADA-ETPAELLPRAPISGPIAPSTVNNTNAMGARVPVAVPG	1325	
Qy	1113	-----TSSPNDISYVQSEYTVTEEGIAG--VATISGVVTSGLP	1153	
Db	1326	PQIGVHSGAALIGVAPVSRSP--VTHGVVTAAP1TTTPRSNLA5VPAQ5MAQOQTHGAP	1383	

QY 1154 -----IPVSHSPVLSVSSITTPAVSISTTSPLQVPTSTSEIVSSNALP 1203
 DB 1384 AGSGNASISRAPSTAPAPTAPTSVPPVS-----STVRPMSTVPTGPPGLAPPA--- 1433
 QY 1204 STVSATASAGSSTATPGPKPAVVSQAAGSTTGATLTSVSTTSPSTASQSLQ 1263
 DB 1434 -----SSGASAGTARANASTMTAPTSQAAMTSANPQSVSPRPS---LTGQGVPA 1485
 QY 1264 SSSSTPTTAEVTV-VSAHSLDKTSHSSTTGALFSLAPSSSSSPGAG-VSSTISQPGI 1321
 DB 1486 IAAASTSRPASGVNPPASSIAPSTK-----SMESAVPTTASGAVSSTVSS--- 1532
 QY 1322 HPLVTSVASTPILPQAAGPTSTPLPQVSPPLVQPVANVPAVQQL-THSQPQ- 1377
 DB 1533 -----LAAATLPSPA--PRYGSNNATPTTPATAPPAAPLPAASVSAPVTCQPS 1582
 QY 1378 -PALPNQPTHCPEDVSDTQPKAPGIDIKITLSEKRLSESHSSSGAQAHSVLSLS 1436
 DB 1583 APASVTPPP-----PSKLLTSTTCLPPTNPLG-----TVA 1476
 QY 1437 VIRSTV--TP--GIPTTAVA-----PSKLLTSTTCLPPTNPLG-----TVA 1476
 DB 1611 AISSIGATPAASIPSSAPALAPVTVPVQQAASAAARLPVTPAPAMHTIAGSVAPQ 1670
 QY 1477 LPTV--PVTPPGVSTPTSTTSQVKGCTAPSKP-----PLTKAPVLPVTELPAGTL 1528
 DB 1671 RPTQSEVQSVAGVHTSSQASTRAHP-VAOSVPRPNSNPTSAAPPAVAGTQVAP 1729
 QY 1529 SEOLPPPPGSLTQSQPLEDLDAQLRKLSPKEXI-----TVAAGVPSMAPT 1578
 DB 1730 VTQ--PAPHRAISVSQSL-----POSVHAAQQAAGTQSHASRPVQGV 1775
 QY 1579 AITBAGTQ--PQKGVSV-----KEGPLATSSGAGVKGMRFOVSVAADQAQEGKRS 1632
 DB 1776 SVQATQAVPRBSTALTPTAQPQVSPVAVSGSV-----PAPSAQSVVA--PA 1823
 QY 1633 DAVSHHESSTSSSVSSPESTLVKPEPNTITIGISSDVEASHKTTASAKSDTG 1692
 DB 1824 PVSTTPVPAATVAPASTVAAAPFTPTRYTAAPALSAATNPAPVPSQHQHITGA 1883
 QY 1693 QPTKGRFOVTTANKVGRFSVSKTEDKITDTYKKEP--VASPPFMDLEQAVLPAV 1750
 DB 1884 RPPQAQAAPATPT-----ITSAPRPPTLAPPP-----PPPPPTPE 1921
 QY 1751 EKPELSEPHLNGSPSPPEAFLSRDVG 1781
 DB 1922 DPPPPPPPAEAPPPEPTPLMPTSSAS 1952

RESULT 12

T30826
 nascent polypeptide-associated complex alpha chain, muscle splice form gp220 - mouse
 N:Alternate names: alpha-NAC protein
 C:Species: Mus musculus (house mouse)
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 24-Nov-1999
 C:Accession: T30826
 R:Yocov, M.V.; St-Arnaud, R.
 Genes Dev. 10, 1763-1772, 1996
 A>Title: Differential splicing in of a proline-rich exon converts alphaNAC into a muscle
 A:Reference number: Z20889; MUID:96312450; PMID:8698236
 A:Accession: T30826
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2187 <YOT>
 A:Cross-references: EMBL:U49363; NID:g1666688; PID:g1666689; PIDN:AAH18732.1
 C:Genetics:
 A:Gene: Naca
 A:Map position: 10
 A:Intons: 24/1; 1996/1; 2024/3; 2050/3; 2099/3; 2142/3; 2183/3
 A>Note: differential splicing converts alphaNAC into a tissue-specific DNA-binding activ
 C:Keywords: alternative splicing; DNA binding; transcription factor

Query Match 4.4%; Score 472; DB 2; Length 2187;

Best Local Similarity 21.5%; Pred. No. 7.8e-07;
 Matches 406; Conservative 208; Mismatches 702; Indels 568; Gaps 94;
 QY 533 PEDVAQEMVSGVCE--GDHKTMAKALDRYSLLIKRKEQQLVREBEKKQKSSSLK 590
 DB 340 PEDDKSSAVTNETLSPSSSVNAGTSLSPKASLV-----DKGSNVAIQ 383
 QY 591 QQVEGSSAQ--TGIKLPASTG-----IPTASTSASVSTQVEPEPEADHQOIQYOQ 644
 DB 384 PLVTQVPAQKTKGLKEIPVSCGATHALDNBSALVMAATVPP----- 429
 QY 645 PSISVLDGTVDSCGSSVFTESRVSQQTVSYSQHEQ-----AHSTGTVP-GHLPSTV 698
 DB 430 -----TSSGIVSKDPAAPVTSILVPP-----AAHKQFPAAPASATLGVVPSPLPATE 476
 QY 699 QASQPHGVYPPSVAVAGSQSQGQSSSILTVSSSQPIQHPOQOQIQOTAPQVQVYS 758
 DB 477 GLKNLP-----ISALVNGAPVSPAQGL-----PTKKDTTLOPLAPALKESSSOS 524
 QY 759 LSQTSSEATTAQAPVOPAPQVLPQVAGKOSTQGVQVAPAEVANAQOPATPTTL 818
 DB 525 ASLEVLSEDTYTKTTGGPAPVYRPAL-AGVATTISLADSP--PAVIRADSCVSPNTV 581
 QY 819 ASSVDSASHDVAGSGSDGNNVPSSGRHGGRTTKRHKVRSRSHKTSRPKRLILN 878
 DB 582 SPLKRSVTDPPMAPAPRTAKNTAPS-----TSP--LVP 612
 QY 879 VSNKGRVVECC--LEHNKRMTVEFKFDLDGNDPEI-----ATIMVNDPILATERES 930
 DB 613 LASEGCPVASSMALSPONASVSETALALSPETIKSVFPDPPLAELSFST-----AR 664
 QY 931 FVDQVEIIEKADEMLSE--DVSVEPEG-----DQGLE-SLQKDDYGFSGQKLEGEFK 982
 DB 665 KUNAVHMESSGSRQGHDPASVTAGTVVCLADSLDTSVSAKSGSALGASAP--L 720
 QY 983 QPIPASSMPQ--QIGITPSSLTQVHAGRRITVSPVPSRKLRESVFPFSEITDYAAS 1039
 DB 721 YPLEVFLPAGLAVOGPKSLNK-----LSFTPPS--SKGAP--VPS 759
 QY 1040 TASP-----GNNLHSAHSSLSLQAASEL-----RAQMEGNTAPPNPNSHT 1083
 DB 760 TGAPESEKAPVPTESISSKQVPAELPSPQKTEPVTASRLISAVQSK-VDPIMSDV 818
 QY 1084 GPTFP-----VPPFLSSIAGVPTTAAT--APVATSSPPNDISTVQSEVTVTEE 1135
 DB 819 TPTSPKTSATAPKXDSATLSLKSVAVATSLSPKAPVAPSE--ATVPTETLPTSLKN 876
 QY 1136 GIAGVATSTGVVTSGLPIPPVSESPVLSVSSITL--PAVSIKTS----- 1182
 DB 877 ALAAATPKETLATS-----IPKVT-SPSPQKTPKSVSLKGAPAMTSKKATEIAASKDVSPS 931
 QY 1183 -----PSLQ--VPTSTSEIVSST--ALY-----PSVTSAT-----SASAG 1216
 DB 932 QPPEVPLLHVPPTPSPKSPVSDTLGALLTSPPKPPATIAETPTTPPKSPKPAASK 991
 QY 1217 STATPGPK-----PP-----AVVQAAGSTVGATLTSVST-----TSPFP 1253
 DB 992 TPATPPEGTAVPVELEPPCKKAPKTAAPKESASATSSKRAPKTAVSKEIPSGVAVP 1051
 QY 1254 STASQSLTQSSSTSTYTLAETVAVVSAHSLDKTSHSST--GLAF----- 1286
 DB 1052 ---LEISLPKETSSTKATPGKSSASPKRSTKATGKPTPGGVAVPVEISLPPKETPQ 1108
 QY 1297 -----SLSAPSSSSSGAGVSVYSIQPGGL--HPLVPS----- 1328
 DB 1109 NATPNESIAASSQKRSKPTSVPKK-TPPGGVTAAPLEIPSPAPQKAPTAVPKQIPTPEDA 1167
 QY 1329 --VIATPILPQAAGPTSTPL-LPQVPSIPLVQPVANVPAVQQLTHSQPQ--PALIP 1382
 DB 1168 VTIAGSLSPKSKAKTAAPKAPATPSGVIAVSGEISPPKTKSTAAPEKENSATLPP 1227
 QY 1383 -NPHTHCPEDVSDTQPKAPGIDIKITLSEKRLSESHSSSGAQAHSVLSLS----- 1436

Db 1228 KSPKTAAPK-----ETPA-----TSSEGVTAIVSEISPSPTTSPACKGVPTLTPKGA 1275
 QY 1437 ---VIESVTP-GLITTAIVADSKLITSTSTCLPTTNPLGVALPVTPVTPGVSTP- 1491
 Db 1276 PNAALSPSPKPKVPAKTA-ABEE-----TSTPSPQKIP-KVAGPKASATPSPSKTPK 1327
 QY 1492 --VSTTSGVPGT-----ASKEPLTKAPVLPVGTGLPA-----GLPSEQLPFP 1535
 Db 1328 TAVKEITSAPSEGVTAIVLEIPSPRPAKPAKTAAP-----KETPAPSPGATTAIVQIPPS 1382
 QY 1536 P--GPSLITSSQOPLLEDLAQRLTISPEXITVTSVAVGVSS-----MAAP--TATTEAGT 1585
 Db 1383 PRKSKKAGSK-----EMPTTPSPBGVTAALPILPISKTKSKASAPETLITVTS- 1432
 QY 1586 QPQGVGVQVEGVPVLAITSSGAGVFKMGKGFVSVADGAKGKNSKEDAKSVHSESSTSE 1645
 Db 1433 --SKMLQGT-VGPKEITLKGATA-----VPLEIPSPHKAP-KITVDPKQVPLITPSPKD 1481
 QY 1646 SSVLSSSSPES-----TLVKEENGITIPGISDVDESANKTTASR-AKSDTGQPTKV 1697
 Db 1482 APITTLASPSPPKAPKPAKTAAPPSRVTTPPEKPAETQKASGTTASKVPAETQEVAVS 1541
 QY 1698 GREQVTTTA-----NKVGRFSVSKTEDIPTDKKEGVASPPFMDLQVLAIVIPKKEK 1752
 Db 1542 SRETPVTPAVPVKNPSHKKTKTIE-----LKEAATLTP-----SPYSPKIPSKKA 1592
 QY 1753 PELSEPHL-NGRPSDEPAFLSRDVDGSGSPHSPPHOLSKSLPQNL----- 1800
 Db 1593 PRTSAPKEFPAPSPKIKPVTTSLAQ-----APPSLQKAPSTTIPKENLAAPVLPSSK 1646
 QY 1801 -----SOSIS-----NSFNSYMSNSNEDIEDEDL-KLELRRLDKHL 1838
 Db 1647 SPAPPARASASLSPATAPOKAPKEATTPSCKKAATEPTIETSTASLEGAPKETSET 1706
 QY 1839 KEIIDLQSRQKHIESLYTKLGKVPVAVIP--PAAPLSGRRRPTKSGSSSSSS-- 1893
 Db 1707 SVSKVLMSPPKASS--SKRASLTLPATLPSLKEASVLS-----PLATSSGDSHISPV 1759
 QY 1894 ----SICNKSPPQLSGNL-----SGGSAASVLRPOOTLHPGNIIPSSGQ 1933
 Db 1760 SDACSTGTTTPQASSEKLPKKGPTAFTFEMLAAPAPSLALATAPIDQ-----KSPGANS 1813
 QY 1934 QLLQPLKPPSSSNIVSAFTSDAISVPSLSABGQGTSSNTVGTATVNSQAQAQPPAMT 1993
 Db 1814 SASPPKCPDSSKKDTKGLPS--AVALAPQTP-----VEKDTSKAIEITLIV 1858
 QY 1994 SSRKGTFTDHLKLVNMAADAMNLSGRSGSGHMYEGBGMARFSAPOGLCISMTSNL 2053
 Db 1859 SPAKG--SDCLH-----SPKG----- 1872
 QY 2054 GGSAPLSAASATSLGHFTKSMCPPOQYGPATPFGAQMGS-TGGPAP-QPLGQFQVGT 2111
 Db 1873 ---PVGSQVATPLAFTSDKVPPEAVSASVAKPAPASLTPLAPVPLPPKQL----- 1925
 QY 2112 SLQNFVITSLQKSISNPPGSLNLT 2135
 Db 1926 -----LESAPGSVLES 1936

RESULT 13
 T18611
 probable serine/threonine-specific protein kinase (EC 2.7.1.-), long splice form - Caeno
 N:Contains: probable serine/threonine kinase, short splice form
 C:Species: Caenorhabditis elegans
 C:date: 15-Oct-1999 #sequence revision 15-Oct-1999 #ext_change 04-Mar-2000
 C:Accession: T18611; T18610; T23144; T23143
 R:McMurray, A.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z18997
 A:Accession: T18611
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1192 <M114>

A:Cross-references: EMBL:Z81027; PIDN:CA854178.1; GSPDB:GN00023; CESP:H39E23.1a
 A:Experimental source: clone AH10
 A:Accession: T18610
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-487,536-1192 <M114>
 A:Cross-references: EMBL:Z81027; PIDN:CA854178.1; GSPDB:GN00023; CESP:H39E23.1b
 A:Experimental source: clone AH10
 R:McMurray, A.
 submitted to the EMBL Data Library, June 1997
 A:Reference number: Z19696
 A:Accession: T23144
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1192 <M114>
 A:Cross-references: EMBL:Z96102; PIDN:CA854262.1; GSPDB:GN00023; CESP:H39E23.1a
 A:Experimental source: clone H39E23
 A:Accession: T23143
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-487,536-1192 <M114>
 A:Cross-references: EMBL:Z96102; PIDN:CA854262.1; GSPDB:GN00023; CESP:H39E23.1b
 A:Experimental source: clone H39E23
 A:Gene: CESP:H39E23.1a; CESP:H39E23.1b
 A:Map position: 5
 A:Insertions: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 487/3; 535/3; 631/1; 825/2; 914/3; 992/2
 A:Keywords: alternative splicing; ATP; phosphotransferase; serine/threonine-specific
 F,1-1192/Product: probable serine/threonine-specific protein kinase, long splice form #s
 F,1-487,536-1192/Product: probable serine/threonine-specific protein kinase, short splice

Query Match 43%; Score 468; DB 2; Length 1192;
 Best local similarity 21.6%; Pred. No. 5,4e-07;
 Matches 274; Conservative 168; Mismatches 508; Indels 316; Gaps 49;

QY 104 QPSIPAAVQSAPEPHEPHEVTAATATQVAAQPPAAAPGEQAVGAPASTVPSSTSKD 163
 Db 71 RPESTSTSSSHRRRAQEMNGSTAT-----ATGGGTGATTTASG 114
 QY 164 RPYSPSLVSGSKEPPPARSGSGGSAKAPQERRSQODIELETKAVGMS----- 215
 Db 115 APA-----ASSGSSARVYSSGSRSHPTSGSSSHARSATGSGMSRSAA 158
 QY 216 --ND-----GRFLKPDIEIGSGFKYVKGDLTEFTVEVAMCELQKRLTKSRQPFKE 268
 Db 159 RRNDQVHVKY-KLTKTGKGNFAVKLAKHVITGHEVAIKI IDKTA LNPSSLQKLPFE 217
 QY 269 AEMLKGLQHPNIVRFYDSWESTVKGKCIVLVTELMTSGTLKTYLKRFPYMKIKVLRSMC 328
 Db 218 VKIMQDLHPNIVKLYQVET-----EQTLVLEVAASGEVADYLVAHGRMKEKARAF 273
 QY 329 RQLKGLQFLHRTPTPIHRLDKCDNIFITGPTGSVKIGDLGLA-TLKRA SFAKSVTGP 387
 Db 274 RQIVSAVOYLHGKN--IHRDLKAENLLDDQM-NIKIADFGFSNFTSLGNKLDFFCGSP 330
 QY 388 FEMAEPMYB-EKYD-ESYDVVAFGMCMEAMSEVPSYSCQAAQIYRVTSGVPAASD 445
 Db 331 PYAAPELBSGKTYDGEVDVMSLVGLVTVLSGLPF-DGQMLKELREVRVNG----- 382
 QY 446 KVAIP-----EVKELIEGIRQNKDERYSIKDL-----LNHAFPOEETVRVLAEBDDG 495
 Db 383 KVRIFPMSTDCENLLKFLVINVQRSSLDNIMDMRMNVGVEDEDLKPFLEPRKDQD 442
 QY 496 EKIAIKMLRLEDIKKLGKGYKDNFAIESFPLERDEVDVAQEVESGYVCEGDHK-- 552
 Db 443 EQ-----RIEKLQIFOLGFNKAIIIESEVEK-----FEDIHATYLLGGRKSDM 488
 QY 553 -----TMAKAIKDRVSL-IKRRRQRO--LVREGEKKKQEESSLK-QVHQSASQSG 602
 Db 489 DASETTMQSLSHSINWSSSLGHPAGVITREHVTSSSASGSASPSRRSRSSATVYG 548
 QY 603 IKOLPASTGTPLASTTSASVSTOVEPEPADHQLOQYOOPISVSLSDGTVDSGQSS 662


```

Db      2152 TGDNTSTPTSTSLASVSTAP-----BGTSAVAVK-LSSLSPVSGPS 2197
QY      817 TLASSVDSASDVASGMSDGNENVPSSGRHGRRTTKRHYKVSAREHEKTSRPKRI 876
Db      2198 TKTFPATESSTVQASSETSSGT-----SVKS-----TSEP----- 2226
QY      877 LNVSKKGRVVECCQLETHNRKKVTFKPLDGDNPETIATIMVNDPILAIERESFVQVR 936
Db      2227 -----ESH-----VT-KLSITSSNPSS-SVPTSP----- 2249
QY      937 EIEKADEMLSEDVVEPEGQGLSLQKDDYGGSGQKLEGEKQIPASMSMQOICI 996
Db      2250 KSTVPSTEQPTSTTSG-QSLTPM-----SNEVLTTSEPHVLSLSLSD--V 2238
QY      997 PTSSLTQVHAGRRFIVSPVESRLRESKVPSEIT-----DTVAASTAQSPGMNLSH 1050
Db      2299 SQSSTTP-----NNLSSTVETPTKTSSEVSLNSHSEPTTEPTLSPDI-LST 2346
QY      1051 SASLSLQQAFASELRRAOM-----TEGNTAP--PNEHRTGPTFPVPV----- 1091
Db      2347 TTNMLSQSSTVSTDRSEISENSEKPTSADELVTSSVTHVASSSDVPTSESPDDLTG 2406
QY      1092 -----FPLSS--IAGVP--TTAAATAPVATSSPNDIST--SVIQSEVTVPTREGI 1137
Db      2407 SSTEINIBASKQKTIISTPTPTDITTAEEPTKSTMSGP-DLSTTSNVLSESTPSESSK 2465
QY      1138 AGVATST--GVVTSGLI-----PIP-----PVSESP 1161
Db      2466 SPVSSSTEGISVVTSTEFKVPSESTISSVLEEDITKTPPSLLEETTTASTSPLEDS 2555
QY      1162 VLSSV-----VSSITTPAVVSIISTSPSLQVPTSEIIVSSTALYPSVATASAS-- 1213
Db      2526 LTVSVRIHELTTSENVPKSESESTTTSESSKPEQBPAGILTSVVVPTSSVLTASEI 2585
QY      1214 -----AGGSTATPGRK-----PPAVSQQAAGSTVGAULTSVSTTSPST 1255
Db      2586 EALTSNTPFOGRPTITTSKSLVKTSTSPSTVSSPSSTKRTTVSTVSTTPTHEET 2645
QY      1256 ASQSLISQSSSTPTPLAETVVAHSLDKTSHSTGLAFSLAPSSSSSPG--AGV 1311
Db      2646 TTSSELTTLTAPSKPT-----ESTTESSEPTTPATSEKPNVSTSRKSTENV 2696
QY      1312 SSVISQPGHPLVIVSVIASTPILPQAGPTSTPLLPQVPSLPPLVQPVANVAVOQLT 1371
Db      2697 ETSTSSQSGSL-----SSTWSTSS-----SEPTNAPAV--T 2726
QY      1372 IHSQPOPALLPQPHTCPEVSDTQPKAPGIDDKLLEKASLSEHSSGA----- 1425
Db      2727 VSSEASTTLEENSTSSP-----TSEA-----SVKLASLTPESITSEAVTVSSR 2772
QY      1426 -----QHASVSL-----ETSLVISTVTPGIPPTVAVAPSK--LITSTSTCLP 1467
Db      2773 APAEITWSSSHREISTVSESEPEPELPLSTVSPNVATASLSPSEPIILSTSSSTPR 2832
QY      1468 TNPPLGT-----VALP-----VTPVVTGQVST--PVSTTSGV 1499
Db      2833 VRLITGPDLLIVSVVPSHGRRONITASSVPSNSTSPILIPSESLTTPQPEPTTTTA 2892
QY      1500 KPGTASRK--PPLTXAPVLPGTELPAGLTPSEQLPPFQPSLTOQOQPEDDAOLRT 1557
Db      2893 KPAITSGKGFPSIQPAEMFTTPAP--PPSNGVGEETNOEBE--CVTST 2940
QY      1558 LSEKXITTVSAVGPVMAAPALTEAGTOPOKGVSOVKEGPVLATSSGAGVFMGRFQV- 1616
Db      2941 TTTEARSLCSTV-----TCHSLATCGSTGVCICRGGFT-----GGTTACSKSKSTA 2987
QY      1617 -----SVAADGAQKGGKKSSEDAKSVHESSTSESSVLSSSPESLTVK--EPNGI- 1666
Db      2988 DCISLPSLQADRAKKNSTRSCGC-----DAGYIGDYGVCSRHPQCVCVRLDNLCSPEANC 3042
QY      1667 -----TIPGISDVE--SAHKTTASEAKSD 1690

```

```

Db      3043 QNRRCQCLPGFTGDGVKCVSIIHERASNCSCD 3074
RESULT 15
S48478
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - Yeast (Saccharomyces cerevisiae)
N/Alternate names: extracellular glucanase; glucan-1,4-galactosylase; protein YJR019C
C/Species: Saccharomyces cerevisiae
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 12-Nov-1999
R/Accession: S48478; E26877; S27281; J06123
R/Rowley, K.
submitted to the EMBL Data Library, October 1994
A/Reference number: S48478
A/Accession: S48478
A/Molecule type: DNA
A/Residues: 1-1367 <ROW>
A/Cross-references: GB:247047; EMBL:238061; NID:G603997; PID:G763364; GSPDB:GNO0009; MIM:
R/Yamashita, I.; Nakamura, M.; Fukui, S.
J. Bacteriol. 169, 2142-2149, 1987
A/Title: Gene fusion is a possible mechanism underlying the evolution of STA1.
A/Reference number: A91831; MIM:87194600; PMID:3106330
A/Accession: A26877
A/Molecule type: DNA
A/Residues: 1-242 <YAM>
A/Cross-references: EMBL:M16164; NID:G172522; PIDN:AAA35014.1; PID:G172525
A/Accession: B26877
A/Molecule type: DNA
A/Residues: 762-1331 <YAN>
A/Cross-references: EMBL:M16165; NID:G172523; PIDN:AAA35015.1; PID:G172526
R/Pardo, J.M.; Ianez, E.; Zalacain, M.; Claros, M.G.; Jimenez, A.
FEBS Lett. 239, 179-184, 1988
A/Title: Similar short elements in the 5' regions of the STA2 and SGA genes from Sacchar
A/Reference number: S27281; MIM:89031230; PMID:3141213
A/Accession: S27281
A/Molecule type: DNA
A/Residues: 1-31 <PAR>
A/Cross-references: EMBL:X13857; NID:G4551; PIDN:CAA32069.1; PID:G4552
R/Lambrecht, M.G.; Bauer, F.F.; Marmur, J.; Pectorius, I.S.
Proc. Natl. Acad. Sci. U.S.A. 93, 8419-8424, 1996
A/Title: Muc1, a mucin-like protein that is regulated by Msi10, is critical for pseudohy
A/Reference number: J06123; MIM:96323237; PMID:8710886
A/Accession: J06123
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-1367 <LAM>
A/Cross-references: GB:U30626; NID:G1304386; PIDN:AA49609.1; PID:G1304387
C/Genetics:
A/Gene: SGD:MUC1; STA2; MAL5; DEX2; SGD:S0001458
A/Cross-references: MIPS:YJR019C; SGD:S0001458
A/Map position: 9R
C/Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase
C/Keywords: glycosidase; hydrolase; polysaccharide degradation; transmembrane protein
F:3-21/Domain: transmembrane #status predicted <TM1>
F:1350-1367/Domain: transmembrane #status predicted <TM2>
Query Match 4.3%; Score 460; DB 1; Length 1367;
Best local Similarity 22.4%; Pred. No. 1,1e-06;
Matches 302; Conservative 164; Mismatches 540; Indels 342; Gaps 54;
QY      585 ESSLKQVQEGSS--ASQGIKQLPSASTGIPASTTAS-----VSTOYEPREPAD 635
Db      229 ESSLTSTSTSESTSTSTSTSESTSTSTTAPATPTTSTCKRKPPTTSTCKEPTTP 288
QY      636 QHQOQLOQOPSLISVSDGTVDGSGQSSVFTSESVSSQOQTVSGSGHQHAGHSTGVGHP 695
Db      289 HHD-----TTPCTKKKTTTSTKCT-----KTTTTPVPTPS 319
QY      696 STVOAQSOPHGVYPPSSVVAQSQSQSQSQSSSLTGVSSSQPIQHPOQOQSIQOTAPPOQTV 755
Db      320 STTBSSAP--VTPSSSTTSSSAPVTSSTT-BSSAPVPTPSSSTTBSSAP----- 370
QY      756 QYSLQSTSTSEATTAQPVSPQAPQVLPVQVSAQKQSTQGVSQ--VAPAEVAVAAQQA 812

```

Db 371 -VTSTTESSAPVTSSTTESSAPVPTSSSTTESSAPVTSSTTESSAPVTSSTES 429
 QY 813 TQPTTLASSVDASHDVASGMSDGN- -NVPSSGHEGRTTKRYKRSVRSRSHK 869
 Db 430 SSAPVTSSTTESSAPVTSSTTESSAPVPTSSS - - - - -ESS 469
 QY 870 SRKMLILNNSNGDRVVECOLETHNRKMTFFKFDLDGDNPELITMWNDFLIAIERE 929
 Db 470 SAP - - - - -VTSTTESSAPVPT - - - - -PSSST - - - - -ESS 496
 QY 930 SFVDQVREIIEKADEMUSEVSVPEEGDQLESLOGKDYFGSGQKLEGEFKOPIPASS 989
 Db 497 S - - - - -APVTSSTTESSAPVPTP - - - - -SSSTTESSAPAPPTSSS 533
 QY 990 MPQIGIPTSLTQVHNSAGRFTVSPVPSRLRESKVP - -SEITD - - - - -TVAAST 1040
 Db 534 TTESSAPVTSSTTESSA - - - - -PVPTSSSTTESSSTPTSTSTTESSAPVPTSSST 588
 QY 1041 AQSPGMNLHSHASSLSIQAFSELRKQMTGEPNTAPNFSHT -GPTFPVPPPL - - - - -S 1095
 Db 589 TESSAPVPTSSSTTESSAPAPPTSSSTTESSAPVTSSTTESSAPVPTSSSTES 648
 QY 1096 SIAGVP - - - - -TAAATAPVPATSSPNDISTSVIOSEVTPTEGJAGVATSGVVTSG 1150
 Db 649 SSAPVPTSSSTTESSAPVPTSSSTTESSAPVTSST - - - - -TESSAPVTSST - - - - -TESS 703
 QY 1151 GLPIPVSESPVLSVVSITTPAVSISTSPLOVPT - - - - -STSEIWSGTALYP 1203
 Db 704 SAPVPTSSSTTES - - - - -SSAPVPTSSSTTESSAPVPTSSSTTESSAPVTSSTES 760
 QY 1204 SVTWSATISAGSGTATPGKPPAVVSOQAAGSTVGATLTSVTSSTSPSTASQLSIOL 1263
 Db 761 SAPVPTSSSTTESSAPVPTSSSTTESSAPVPTSSSTTESVAPVPTSSSNITS 820
 QY 1264 SSSTSTPLAETVVVSHSLDKTSHSSTTGAFSLAPSSSSSPAGVSYISQPGGLHP 1323
 Db 821 SAPSTPSSST - -ESSSVPTPSSSTTE - - - - -SSAPVSSSTTESSVAP - - - - - 865
 QY 1324 LVIPSIASTPILPOAGPTSTPILPOVPSIP - - - - -PLVOQVAN -VPANO -QTLHSGQPPA 1379
 Db 866 - - - - -VPTSSSNITS - - - - -PSSIPSSSTTESPTSTPTVPSKXPGSQETSVSSTETT 921
 QY 1380 LLPNOHHCPEVSDPOPKAPGIDDIKLEKLSLSEHSSGQAQASVLETSVIE 1439
 Db 922 IVPTK - - - - -TTSVTPSTTTTITTVCSGTNSAGETSGCSPKTV - - - - - 963
 QY 1440 STVTPGIPPTVAVPSKLTSTSTCLPPTNLPLGVALPVPVTPQGVST - -PVSTT - - 1495
 Db 964 TTVPTTTTTSVTSTTTTITTVCSGTN - - - - -SAGETTSKSPKTTTTPCSTPS 1018
 QY 1496 -TSGVKPGTAPSKPPLKAPVLPVGTLPAGTLPSEQLP - - - - -PFGPSLTOSQPLE 1548
 Db 1019 ETASESTTSPPTPTVTVSTTVVTEYSTKPGSEITTFVTKNIPTTYLT - - - - - 1071
 QY 1549 DLDQKRLTSPKXTVTSAGPVMAAPTAIT - - - - -EAGT - - - - -QPOKGVSYKEGPVL 1600
 Db 1072 - - - - -TIAP - - - - -TPSVTVTNPTPTTITTVCSGTNSAGETSGCSPKTVTTV 1118
 QY 1601 ATSSGAVFKMGKRGQVSVAA - - - - -DGAQKGNKS -EDAKSVHESSTESSV 1648
 Db 1119 PCSTGTGTYTEATTLVTAVTTVTTESSGTNSAGKTTGYTKSV - - - - -PTTYVT 1174
 QY 1649 LSSSPSTLVKPEPNGITPGIISDVPSAHKTTASLAKS - - - - - 1689
 Db 1175 LAPSAVPTPATNAVPTTITTECSAATNAAGETTSVCSAKTIVSSASAGENTAPGATPV 1234
 QY 1690 DTGQPTKGRPOVTTTANKVGFVSXKEDKT - - - - -DTKKEGVASPPFMDLEQ 1740
 Db 1235 TTAIPT - - - - -TVITTESSVGNAGEITTYGTTKSIPTTYITLIPSGNKANEYIVAT 1289
 QY 1741 AVLPAVIPKKEK - - - - -PELSERSHLNPGSPDPEAFLSRDVGSGSPHSPHQ 1789
 Db 1290 ATNPISIKTISQALTTASASVAVPTTBS -LTGP - - - - -LQSASGS - - - - -A 1331

QY 1790 LSKSLPSQNLQSLNSFNSSYMSGDN 1817
 Db 1332 VATSVP - - - - -SISTYQGAAN 1349

Search completed: September 24, 2004, 01:11:50
 Job time : 80 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:11:09 ; Search time 204 Seconds

(without alignments)
3366,909 Million cell updates/sec

Title: US-10-010-720-14

Perfect score: 10812
Sequence: 1 MSGAAAEKQSTPGSLFLSP.....NISMQKISINSPGSLNRTT 2136

Scoring table:

BLOSUM62

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09D_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10794.5	99.8	2135	US-10-362-892-9	Sequence 9, Appli
2	10794.5	99.8	2135	US-10-288-798-9	Sequence 9, Appli
3	10677	98.8	2382	US-10-336-472-230	Sequence 230, App
4	10677	98.8	2382	US-10-196-935A-2	Sequence 2, Appli
5	10677	98.8	2382	US-10-052-648A-10	Sequence 40, Appli
6	10677	98.8	2382	US-10-408-765A-1404	Sequence 1404, Ap
7	9204	85.1	2126	US-10-052-648A-39	Sequence 39, Appli
8	3408	31.5	670	US-10-052-648A-38	Sequence 38, Appli
9	2439	22.6	1743	US-10-460-545-2	Sequence 2, Appli
10	2130.5	19.7	1345	US-10-433-794-17	Sequence 17, Appli
11	2126	19.7	1231	US-10-052-648A-37	Sequence 37, Appli
12	2126	19.7	1231	US-10-196-935A-4	Sequence 4, Appli
13	2042.5	18.9	1251	US-09-925-301-1286	Sequence 1286, Ap
14	2042.5	18.9	1251	US-10-114-270-80	Sequence 80, Appli
15	2038.5	18.9	1084	US-10-415-011-17	Sequence 17, Appli

16	2038.5	18.9	1234	US-10-052-648A-36	Sequence 36, Appli
17	1959.5	18.1	560	US-10-114-693-84	Sequence 84, Appli
18	1866.5	17.3	1356	US-10-114-270-78	Sequence 78, Appli
19	1861	17.2	359	US-10-664-421-152	Sequence 152, App
20	1804.5	16.7	779	US-10-353-929-49	Sequence 49, Appli
21	1776	16.4	1069	US-10-182-243-50	Sequence 50, Appli
22	1677	15.5	324	US-09-933-767-533	Sequence 533, App
23	1677	15.5	324	US-10-004-860-533	Sequence 533, App
24	1677	15.5	324	US-10-023-282-533	Sequence 533, App
25	1677	15.5	324	US-10-023-282-533	Sequence 533, App
26	1572.5	14.5	557	US-10-052-648A-12	Sequence 12, Appli
27	1362.5	12.6	663	US-10-108-260A-3492	Sequence 3492, Ap
28	1323	12.2	309	US-09-862-027-37	Sequence 37, Appli
29	1280.5	11.8	1381	US-10-233-131-25	Sequence 25, Appli
30	1280.5	11.8	1381	US-10-233-131-25	Sequence 34, Appli
31	761.5	7.0	613	US-10-240-145-77	Sequence 77, Appli
32	745.5	6.9	516	US-09-862-027-39	Sequence 39, Appli
33	745	6.9	587	US-10-437-963-175930	Sequence 175930, A
34	736	6.8	567	US-10-424-599-222750	Sequence 222750, A
35	729	6.7	567	US-10-425-114-69051	Sequence 69051, A
36	723.5	6.7	618	US-10-437-963-164986	Sequence 164986, A
37	722.5	6.7	746	US-10-425-114-72414	Sequence 72414, A
38	706	6.5	604	US-10-424-599-261688	Sequence 261688, A
39	702.5	6.5	549	US-10-425-114-64415	Sequence 64415, A
40	694.5	6.4	477	US-10-437-963-144326	Sequence 144326, A
41	688	6.4	412	US-10-424-599-170860	Sequence 170860, A
42	683	6.3	658	US-10-437-963-183335	Sequence 183335, A
43	681	6.3	864	US-10-437-963-183335	Sequence 38, Appli
44	678.5	6.3	677	US-09-862-027-38	Sequence 183311, A
45	678.5	6.3	678	US-10-437-963-183341	

ALIGNMENTS

```

RESULT 1
US-10-362-892-9
; Sequence 9, Application US/10362892
; Publication No. US20040038861A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; BANDMAN, Olga
; APPLICANT: NGUYEN, Danielle B.; WALIA, Narinder K.
; APPLICANT: HAFALIA, April J.A.; YAO, Monique G.
; APPLICANT: GANDHI, Ameena R.; GURRATAN, Rajagopal
; APPLICANT: DING, Li; PATTERSON, Chandra S.
; APPLICANT: YEH, Henry; BAUGHN, Mariah R.
; APPLICANT: TRIBOULEY, Catherine M.; THORNTON, Michael B.
; APPLICANT: ELLIOTT, Vicki S.; LU, Yan
; APPLICANT: ISON, Craig H.; AU-YOUNG, Janice K.
; APPLICANT: TANG, Y. Tom; AZIMZAI, Valda
; APPLICANT: BURRILL, John D.; MARCUS, Gregory A.
; APPLICANT: ZINGLER, Kurt A.; LU, Dying Aina M.
; APPLICANT: LAU, Preeti G.; RAMKOWAR, Jayalakshmi
; APPLICANT: WARREN, Bridget A.; KEARNEY, Liam
; APPLICANT: POLICKY, Jennifer L.; THANGAVELU, Kavitha
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PF-0209 USN
; CURRENT APPLICATION NUMBER: US/10362, 892
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229, 873
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/231, 357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232, 654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/234, 902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/236, 499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238, 389

```



```

PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/240,542
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 2135
TYPE: PR
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US2004003881A1 5502218CD1
US-10-362-892-9

Query Match 99.8%; Score 10794.5; DB 12; Length 2135;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2134; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MSGGAAGKOSTPGSLFLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEYRRRHT 60
DB 1 MSGGAAGKOSTPGSLFLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEYRRRHT 60
QY 61 MDKDSRGAATTTTTERFRFRSVICDSNATALELPGILPLDPPSIPAAVPOGAPPEPH 120
DB 61 MDKDSRGAATTTTTERFRFRSVICDSNATALELPGILPLDPPSIPAAVPOGAPPEPH 120
QY 121 REETVATATSQVAQOPPAAPAEQAVAGPAPSTVSSSTSKRPVQPSLVGSKKEPPP 180
DB 121 REETVATATSQVAQOPPAAPAEQAVAGPAPSTVSSSTSKRPVQPSLVGSKKEPPP 180
QY 181 ARSSGGGSKKEPQERSQOQDDIEELETKAVGMSNGRFLKPIELIGRGSFKTVYGLD 240
DB 181 ARSSGGGSKKEPQERSQOQDDIEELETKAVGMSNGRFLKPIELIGRGSFKTVYGLD 240
QY 241 TETVEVAMCLOQRKLTSEKQRFKEAEMLKGLQHPNIVRFYDSVESTVYKGCIVLV 300
DB 241 TETVEVAMCLOQRKLTSEKQRFKEAEMLKGLQHPNIVRFYDSVESTVYKGCIVLV 300
QY 301 TELMTSGTLTKYKRFKVMKIVLRSCROILKGLQFLHTTTPPIIHRDLKCDNIFFITGP 360
DB 301 TELMTSGTLTKYKRFKVMKIVLRSCROILKGLQFLHTTTPPIIHRDLKCDNIFFITGP 360
QY 361 TGSVKIGLGLATLKRAFAKSVIGTPEFMAPEMEEKIDESVDVYAFGCMLEMAISEY 420
DB 361 TGSVKIGLGLATLKRAFAKSVIGTPEFMAPEMEEKIDESVDVYAFGCMLEMAISEY 420
QY 421 PYSECQNAAQYRRVTSGVKPASFQKVAIPVKEIIIEGCIRONKDERYSIKDNLNHAFFQ 480
DB 421 PYSECQNAAQYRRVTSGVKPASFQKVAIPVKEIIIEGCIRONKDERYSIKDNLNHAFFQ 480
QY 481 BETGVVLAEEDEGEKIAIKLWLRIEDIKLKGKYKONEMAI EFSFULERVPEVDVQEM 540
DB 481 BETGVVLAEEDEGEKIAIKLWLRIEDIKLKGKYKONEMAI EFSFULERVPEVDVQEM 540
QY 541 VESGVVCEGDHKTAKAIKDRVSLIKRKROQLVREOEKKOESESLKQOVESSASQ 600
DB 541 VESGVVCEGDHKTAKAIKDRVSLIKRKROQLVREOEKKOESESLKQOVESSASQ 600
QY 601 TGIKOLPASSTGIPASTTSASVSTOVEPEBEADQHOLOLQYQPSISVLSGDIVDSQGS 660
DB 601 TGIKOLPASSTGIPASTTSASVSTOVEPEBEADQHOLOLQYQPSISVLSGDIVDSQGS 660
QY 661 SSVFTESRNVSOQTVSYSGQHEQAHSTGTVPCHITSTVQAGSQPHGVYPPSSVAGQSQGS 720
DB 661 SSVFTESRNVSOQTVSYSGQHEQAHSTGTVPCHITSTVQAGSQPHGVYPPSSVAGQSQGS 720
QY 721 QPSSSSTLGVSSSQIOPHPOQOOGIQQTAPPOQTVQYSLQSTSTSEATTPAPVQOPAP 780
DB 721 QPSSSSTLGVSSSQIOPHPOQOOGIQQTAPPOQTVQYSLQSTSTSEATTPAPVQOPAP 780
QY 781 QVLPQVSAKQSTQGVSVQAPAEPAVAQAPQATQPTTLASSVDASHSDVASMSDGENEV 840
DB 781 QVLPQVSAKQSTQGVSVQAPAEPAVAQAPQATQPTTLASSVDASHSDVASMSDGENEV 840

```

```

QY 841 PSSSGRHEGRTTKRHRKSVRSRREKTSRPELRIINVSNGDRVVECOLETHNKKMT 900
DB 840 PSSSGRHEGRTTKRHRKSVRSRREKTSRPELRIINVSNGDRVVECOLETHNKKMT 900
QY 901 FKFDLDGNDPEELATIMVNDPILAIERESFVDQVEIIEKADENLSEDSVEBEDQGL 960
DB 900 FKFDLDGNDPEELATIMVNDPILAIERESFVDQVEIIEKADENLSEDSVEBEDQGL 960
QY 961 ESTQKDDYGFSSQKLEGEFKQPIPASSMPQOIGIPSSLTQVHSAGRRFVSPVES 1020
DB 960 ESTQKDDYGFSSQKLEGEFKQPIPASSMPQOIGIPSSLTQVHSAGRRFVSPVES 1020
QY 1021 RLRESKVPSEITDQVAASTAGSPGMNLSHASLSLQCAFSELRRAQTEGNTAPPNF 1080
DB 1020 RLRESKVPSEITDQVAASTAGSPGMNLSHASLSLQCAFSELRRAQTEGNTAPPNF 1080
QY 1081 SHTGPTFPVVPPLSSIAGVPTTAATAAPVPATSSPPNDISTVYQSEVTVPIREGIAGV 1140
DB 1080 SHTGPTFPVVPPLSSIAGVPTTAATAAPVPATSSPPNDISTVYQSEVTVPIREGIAGV 1140
QY 1141 ATSTGVTSGGLPIPPVSESPVLSVSVSTTIPAVVISITSPSLQVPTSETIWSSTA 1200
DB 1140 ATSTGVTSGGLPIPPVSESPVLSVSVSTTIPAVVISITSPSLQVPTSETIWSSTA 1200
QY 1201 LYSVTVSATSASAGSGTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSPBSTAQLS 1260
DB 1200 LYSVTVSATSASAGSGTATPGKPPAVVSOQAAGSTTVGATLTSVSTTSPBSTAQLS 1260
QY 1261 IOLSSSTPTPLAEVTVVVSASHSDKTSHSSSTGLAFSLAPSSSSSAGVSVSTISQPGG 1320
DB 1260 IOLSSSTPTPLAEVTVVVSASHSDKTSHSSSTGLAFSLAPSSSSSAGVSVSTISQPGG 1320
QY 1321 LHPVTVSVASAPPILOAAGPSTPLPQVPSIPPIVOPVANYPAVQOILIHQOPPAL 1380
DB 1320 LHPVTVSVASAPPILOAAGPSTPLPQVPSIPPIVOPVANYPAVQOILIHQOPPAL 1380
QY 1381 LPNQPTHCEPVDSDTOPKAPGIDIDIKLEBKRLSLSSESSSQAQASVSLSTLVIES 1440
DB 1380 LPNQPTHCEPVDSDTOPKAPGIDIDIKLEBKRLSLSSESSSQAQASVSLSTLVIES 1440
QY 1441 TVTPGIPTTAAVAPSKLITSTSTLCPPTNPLPLGVVALPVTPVTPGVQSVTSTTSQVK 1500
DB 1440 TVTPGIPTTAAVAPSKLITSTSTLCPPTNPLPLGVVALPVTPVTPGVQSVTSTTSQVK 1500
QY 1501 PGTAPSKPPLTKAPVLVGLTELPAGTLPSBQLPFPQPSLQSQOPLBEDDAQRLTISP 1560
DB 1500 PGTAPSKPPLTKAPVLVGLTELPAGTLPSBQLPFPQPSLQSQOPLBEDDAQRLTISP 1560
QY 1561 EXITVTSAGVSMAPALTEAGTOPQKGVSYQVKEGVFLATSSGAGVFXMGRFQVSVA 1620
DB 1560 EXITVTSAGVSMAPALTEAGTOPQKGVSYQVKEGVFLATSSGAGVFXMGRFQVSVA 1620
QY 1621 DGAQKEGKNKSEDAKSVHESSTESSVYSSSSESTLYKREPNGITIPGISDVPEBAH 1680
DB 1620 DGAQKEGKNKSEDAKSVHESSTESSVYSSSSESTLYKREPNGITIPGISDVPEBAH 1680
QY 1681 KTTSEAKSDPDGPTKVGRFVTTTANKVGRFSYKTEDKITTDIKKEGPVAPPEMDLEQ 1740
DB 1680 KTTSEAKSDPDGPTKVGRFVTTTANKVGRFSYKTEDKITTDIKKEGPVAPPEMDLEQ 1740
QY 1741 AVLPAVTPKKEKPELSEPSHLNGPSSDEEAFLSRDYVDGSGPHSPHQLSSKSLPSQNL 1800
DB 1740 AVLPAVTPKKEKPELSEPSHLNGPSSDEEAFLSRDYVDGSGPHSPHQLSSKSLPSQNL 1800
QY 1801 SQSLNSFNSSYMSDNESEDIEDLKLRLRPLKXKLEIOLDSRKHEIESLYTLGL 1860
DB 1800 SQSLNSFNSSYMSDNESEDIEDLKLRLRPLKXKLEIOLDSRKHEIESLYTLGL 1860
QY 1861 KVPAPVITIPPAAPLISGRRRPPTKSGKSSRSSSLGNKSPOLSGNLSCQSAASVTHPQOT 1920
DB 1860 KVPAPVITIPPAAPLISGRRRPPTKSGKSSRSSSLGNKSPOLSGNLSCQSAASVTHPQOT 1920

```

QY 1921 LHPNIPESGONOLLOPKSPSSDNLYSAFTSPGALSVPBLSAPGCGTSTNTVGATV 1980
 DB 1920 LHPNIPESGONOLLOPKSPSSDNLYSAFTSPGALSVPBLSAPGCGTSTNTVGATV 1979
 QY 1981 NSOAAQAPPMATSRKGTFTDDHLKLVNNAARDAMNLSGRGSGHNNYGPGRARFS 2040
 DB 1980 NSOAAQAPPMATSRKGTFTDDHLKLVNNAARDAMNLSGRGSGHNNYGPGRARFS 2039
 QY 2041 APGOLCISMTSNLGGASAPISAASATSLGHFTKSMCPPOOYGFPAITPFGAOWSGTGPAPO 2100
 DB 2040 APGOLCISMTSNLGGASAPISAASATSLGHFTKSMCPPOOYGFPAITPFGAOWSGTGPAPO 2099
 QY 2101 PLGQFPQVGTASLQNFNINSNQKISNPPGSGNLT 2136
 DB 2100 PLGQFPQVGTASLQNFNINSNQKISNPPGSGNLT 2135
 RESULT 2
 US-10-288-798-9
 Sequence 9, Application US/10288798
 Publication No. US20030207299A1
 GENERAL INFORMATION:
 APPLICANT: BANDMAN, Olga; NGUYEN, Daniel B;
 APPLICANT: WALIA, Narinder K.; HARALYA, April J.A.;
 APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
 APPLICANT: GURURAJAN, Rajagopal; DING, Li;
 APPLICANT: PATTERSON, Chandra; YUE, Henry;
 APPLICANT: BAUGHN, Mariah R.; TRIBOLEY, Catherine M.;
 APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
 APPLICANT: LU, Yan; ISON, Craig H.;
 APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
 APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
 APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
 APPLICANT: LU, Dyoung Alina M.; LAI, Preeti G.;
 APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
 APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
 APPLICANT: THANHAVEU, Kavitha; BURFORD, Neil
 TITLE OF INVENTION: HUMAN KINASES
 FILE REFERENCE: PI-0209 USA
 CURRENT APPLICATION NUMBER: US/10/288,798
 CURRENT FILING DATE: 2002-11-01
 PRIOR APPLICATION NUMBER: PCT/US01/27219
 PRIOR FILING DATE: 2001-08-31
 PRIOR APPLICATION NUMBER: US 60/240,542
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: US 60/238,389
 PRIOR FILING DATE: 2000-10-06
 PRIOR APPLICATION NUMBER: US 60/236,499
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: US 60/234,902
 PRIOR FILING DATE: 2000-09-22
 PRIOR APPLICATION NUMBER: US 60/232,654
 PRIOR FILING DATE: 2000-09-14
 PRIOR APPLICATION NUMBER: US 60/231,357
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: US 60/229,873
 PRIOR FILING DATE: 2000-08-31
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: PERL Program
 SEQ ID NO 9
 LENGTH: 2135
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. US20030207299A1 5502218CD1
 US-10-288-798-9
 Query Match 99.8%; Score 10794.5; DB 15; Length 2135;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2134; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 1 MSGGALEKQSTGSLFLSPAPAPKXGSSSDSVGEKLGAAADAVTGRTEYRRRRHT 60

DB 1 MSGGALEKQSTGSLFLSPAPAPKXGSSSDSVGEKLGAAADAVTGRTEYRRRRHT 60
 QY 61 MDKSRGAAATTTTTERHFFRRSVYCDSNATALELPGILPLSPDIPAAVQSAPEPH 120
 DB 61 MDKSRGAAATTTTTERHFFRRSVYCDSNATALELPGILPLSPDIPAAVQSAPEPH 120
 QY 121 REETVATATQVAAQPPAAAPGEQAVAGAPETVSTSTKDDPVQSPSLVGSKEEPP 180
 DB 121 REETVATATQVAAQPPAAAPGEQAVAGAPETVSTSTKDDPVQSPSLVGSKEEPP 180
 QY 181 ARSGGGGSAKEPEEHSQOQDDIELEETAVAGSNDRFLKFDIEIGRGSFKTYVKGLD 240
 DB 181 ARSGGGGSAKEPEEHSQOQDDIELEETAVAGSNDRFLKFDIEIGRGSFKTYVKGLD 240
 QY 241 TETVEAVACELQDRKLTKESERFKEEAEMLKGLQHPNIVRFYDSWESTVYKGCIVLV 300
 DB 241 TETVEAVACELQDRKLTKESERFKEEAEMLKGLQHPNIVRFYDSWESTVYKGCIVLV 300
 QY 301 TELMTSGTLTKYLRKRFVYMKIKVLRSMCRQILKGLQFLHRTPTPIHRDLKCDNIFITGP 360
 DB 301 TELMTSGTLTKYLRKRFVYMKIKVLRSMCRQILKGLQFLHRTPTPIHRDLKCDNIFITGP 360
 QY 361 TGSYKIGDLGLATLKRAFPKSVIGTPEFNAPEMYEEKYDSVYVAFGMCMLEMTSEY 420
 DB 361 TGSYKIGDLGLATLKRAFPKSVIGTPEFNAPEMYEEKYDSVYVAFGMCMLEMTSEY 420
 QY 421 PYSCQAAQIYRRVTSQVYCPASDCKAIPBVEIIEGCIKONKERYSIKDLINHAFO 480
 DB 421 PYSCQAAQIYRRVTSQVYCPASDCKAIPBVEIIEGCIKONKERYSIKDLINHAFO 480
 QY 481 BEGVRYEALBEDGEKIAIKMLRIEDIKKLGXYKDNDAIESFLBEDVPEVDAQEM 540
 DB 481 BEGVRYEALBEDGEKIAIKMLRIEDIKKLGXYKDNDAIESFLBEDVPEVDAQEM 540
 QY 541 VESGYVEGDHKTMAKAIKDRVSLIKKRFQRLVHEEQRKQEEBSLQOYVEQSSASQ 600
 DB 541 VESGYVEGDHKTMAKAIKDRVSLIKKRFQRLVHEEQRKQEEBSLQOYVEQSSASQ 600
 QY 601 TGIKOLPASTGIPIYASTTSASVSTOVEPEPADHQOLQOYQOPSTISVLSDTGTVBSGG 660
 DB 601 TGIKOLPASTGIPIYASTTSASVSTOVEPEPADHQOLQOYQOPSTISVLSDTGTVBSGG 660
 QY 661 SSVFTSRVSSQOTVSGSQHEQAHSTGVPHIPSTVQAOQSPHGVYPPSSVAQGSQG 720
 DB 661 SSVFTSRVSSQOTVSGSQHEQAHSTGVPHIPSTVQAOQSPHGVYPPSSVAQGSQG 720
 QY 721 QPSSSSITGVSSSQPIQHP-0QOGIQTAPQQTGVYSLQSSTSEATTAGVQSPQAP 779
 DB 721 QPSSSSITGVSSSQPIQHP-0QOGIQTAPQQTGVYSLQSSTSEATTAGVQSPQAP 779
 QY 781 QVLPQVSAQKOSTQVSAVAPAPVAVQAOPTPTTLASSVDSAHSDVAGSGMDGNNV 840
 DB 781 QVLPQVSAQKOSTQVSAVAPAPVAVQAOPTPTTLASSVDSAHSDVAGSGMDGNNV 839
 QY 841 PSSSGHREGTTRGHRKSVRSRSHKTSRPLRLTLYNSNKGDRVVECOLETHNKKMT 900
 DB 841 PSSSGHREGTTRGHRKSVRSRSHKTSRPLRLTLYNSNKGDRVVECOLETHNKKMT 899
 QY 901 FKFDLDGNDPEEIAITMANNDFILATERESFVDQVEIEIEKADENMLSEDSVVEPEDQGL 960
 DB 901 FKFDLDGNDPEEIAITMANNDFILATERESFVDQVEIEIEKADENMLSEDSVVEPEDQGL 959
 QY 961 ESIQKDDYGFSSQCLBEFQKPIPASSMPQOIGIPITSLTQVVSAGRRFVSVPEPS 1019
 DB 961 ESIQKDDYGFSSQCLBEFQKPIPASSMPQOIGIPITSLTQVVSAGRRFVSVPEPS 1019
 QY 1021 RLRESKVPFSEITDTVAASTAOSPGMNLGSHSASSLSLQAFSLRAQMTGEGNTAPPNV 1080
 DB 1021 RLRESKVPFSEITDTVAASTAOSPGMNLGSHSASSLSLQAFSLRAQMTGEGNTAPPNV 1079
 QY 1081 SHTGTFPVVPPFLSIAGVPTTAAATAPVATSSPNDISTVIOSEVTVPTGEGIAGV 1140

```

Db      1080 SHTGPTFPVPPPLSSAGVPTTAATAAPVATSPSPNDISTVIOSEVTVPTHEGAGV 1139
QY      1141 ATSTGVTSGLDPIPVSESPVLSSVVSITTIIPAVSISTTSPSLQVPTSTSEIIVSSIA 1200
Db      1140 ATSTGVTSGLDPIPVSESPVLSSVVSITTIIPAVSISTTSPSLQVPTSTSEIIVSSIA 1199
QY      1201 LYSVTVSATSASAGSGTATPPGPPAVVQQAAGSTTVGATLTSTSTTSFPTASQLS 1260
Db      1200 LYSVTVSATSASAGSGTATPPGPPAVVQQAAGSTTVGATLTSTSTTSFPTASQLS 1259
QY      1261 IOLSSSTSTPLAETVAVVASHSLDKTSHSSTTGLAFSLAPSSSSSGAGVSYISQPGG 1320
Db      1260 IOLSSSTSTPLAETVAVVASHSLDKTSHSSTTGLAFSLAPSSSSSGAGVSYISQPGG 1319
QY      1321 LHPVTVSVIASVPILOAAGPTSTPLLPQVPSIPIVQPVANVAVQQLIHQOPAL 1380
Db      1320 LHPVTVSVIASVPILOAAGPTSTPLLPQVPSIPIVQPVANVAVQQLIHQOPAL 1379
QY      1381 LPPPHTHCEPVSDTOPKAPGIDDIKTELEKLSLSEHSSSGAQAHSVLETSVIES 1440
Db      1380 LPPPHTHCEPVSDTOPKAPGIDDIKTELEKLSLSEHSSSGAQAHSVLETSVIES 1439
QY      1441 TVTPTGIPITVAAPSKLTSTSTSTCLPPTNLPLGTVALPVTVPVTPGVVSIPVSTTSVX 1500
Db      1440 TVTPTGIPITVAAPSKLTSTSTSTCLPPTNLPLGTVALPVTVPVTPGVVSIPVSTTSVX 1499
QY      1501 PGTAPSKPPLTKAPVLPVGTTELPAGTLPSEQLPFPFGBSLTQSQQPLEDDAQLRRLTSP 1560
Db      1500 PGTAPSKPPLTKAPVLPVGTTELPAGTLPSEQLPFPFGBSLTQSQQPLEDDAQLRRLTSP 1559
QY      1561 EXTIVTSVAVGVSAAPATATEACTOPKGVQVKEGVLTATSSGAGVFKMGKQVAVAA 1620
Db      1560 EXTIVTSVAVGVSAAPATATEACTOPKGVQVKEGVLTATSSGAGVFKMGKQVAVAA 1619
QY      1621 DGAQKEGKNSEDAKSVHFESESTSESSVLSSSSESTLVKEPEPGITIPGISDVPESAH 1680
Db      1620 DGAQKEGKNSEDAKSVHFESESTSESSVLSSSSESTLVKEPEPGITIPGISDVPESAH 1679
QY      1681 KTTASEAKSDTGQPTKRGFOVTTTANKVGRFVSQKTEBDKTTDKKGGPVASPPFMLEQ 1740
Db      1680 KTTASEAKSDTGQPTKRGFOVTTTANKVGRFVSQKTEBDKTTDKKGGPVASPPFMLEQ 1739
QY      1741 AVLPAAVLPKKEKPELSEPSHNGSPSDEPAFLSRVDVDSGSPHSPOLSKSLPSONL 1800
Db      1740 AVLPAAVLPKKEKPELSEPSHNGSPSDEPAFLSRVDVDSGSPHSPOLSKSLPSONL 1799
QY      1801 SOSLSNFSNSSYMSGSDNESDIEDDLKELRLRDKLKE1QDLOSRQKHEIESLYTKLG 1860
Db      1800 SOSLSNFSNSSYMSGSDNESDIEDDLKELRLRDKLKE1QDLOSRQKHEIESLYTKLG 1859
QY      1861 KVPAAVITPPAAPLGGRRRRPTKSGKSSSSLSGNKSPOLSGNLSCQSAASVILHPQOT 1920
Db      1860 KVPAAVITPPAAPLGGRRRRPTKSGKSSSSLSGNKSPOLSGNLSCQSAASVILHPQOT 1919
QY      1921 LHPGNIPESONOLLOPLKPSPSSDNLYSAPTSIGAIVSVLSLAPGGGTSTNTVAGTV 1980
Db      1920 LHPGNIPESONOLLOPLKPSPSSDNLYSAPTSIGAIVSVLSLAPGGGTSTNTVAGTV 1979
QY      1981 NSQAQAOPPAMTSSRKGTFTDHLKLVDNNAARDAMNLSGRRGSGKHNNYEGPGMARFS 2040
Db      1980 NSQAQAOPPAMTSSRKGTFTDHLKLVDNNAARDAMNLSGRRGSGKHNNYEGPGMARFS 2039
QY      2041 APGOLCISMTSNLGSAPISASAATSLGHFTKSMCPPOQYGPATPPFGAOWSGTGAPAQ 2100
Db      2040 APGOLCISMTSNLGSAPISASAATSLGHFTKSMCPPOQYGPATPPFGAOWSGTGAPAQ 2099
QY      2101 PLGGFQPVGTASLQNFNTSNLOKSISNPPGSLRTT 2136
Db      2100 PLGGFQPVGTASLQNFNTSNLOKSISNPPGSLRTT 2135

```

```

; Sequence 230. Application US/10336472
; Publication No. US20040043929A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Ballinger, Robert A.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Chant, John S.
; APPLICANT: Bergns, Constance
; APPLICANT: Ganggoli, Bsha A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Furtak, Valeriy
; APPLICANT: Gilbert, Jennifer A.
; APPLICANT: Gunther, Erik
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Ji, Weizhen
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: Miller, Charles E.
; APPLICANT: Miller, Isabelle
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patrajan, Meera
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R.
; APPLICANT: Mishra, Vishnu
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Splet, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Ort, Tatiana
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Tchenerov, Velizar T.
; APPLICANT: Verne, Corine A.M.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; RTIE REFERENCE: 21402-533C
; CURRENT APPLICATION NUMBER: US/10/336,472
; PRIOR FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: 09/746,491
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 10/005,041
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 10/023,681
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 10/024,212
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 10/055,569
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 10/080,334
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 10/092,900
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 10/136,826
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 10/236,417
; PRIOR FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 60/345,092
; PRIOR FILING DATE: 2002-01-04
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: Cytaseqdist version 0.1
; SEQ ID NO 230
; LENGTH: 2362
; TYPE: PRT

```

ORGANISM: MNKI
US-10-336-472-230

Query Match 98.8%; Score 10677; DB 12; Length 2382;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;

QY 1 MSGGAEEQSSSTPGSLFLSPAPAPKXSSDSSVGEKLGAAADAVTGTTEYYRRRHT 60
DB 1 MSGGAEEQSSSTPGSLFLSPAPAPKXSSDSSVGEKLGAAADAVTGTTEYYRRRHT 60
QY 61 MDKSRGAAATTTTTEHFRFRSVCDSNATALELPGPLSLPQPSITAAAPQSAPEPPH 120
DB 61 MDKSRGAAATTTTTEHFRFRSVCDSNATALELPGPLSLPQPSITAAAPQSAPEPPH 120
QY 121 REETVATATSOVAQOPPAAPGEOAVAGPAPSTPSTSKDRPVSOPLVSGKEPP 180
DB 121 REETVATATSOVAQOPPAAPGEOAVAGPAPSTPSTSKDRPVSOPLVSGKEPP 180
QY 181 ARSGGGGSAKEPOEERSQODDIELETKAVGMSNDGRFLKFDIEIGRSGFKTVYKGLD 240
DB 181 ARSGGGGSAKEPOEERSQODDIELETKAVGMSNDGRFLKFDIEIGRSGFKTVYKGLD 240
QY 241 TETIVEVAMCELODRKLTKESEKPFKEAEMLKGLQHNIVRFYDSWESTYKGCIVLV 300
DB 241 TETIVEVAMCELODRKLTKESEKPFKEAEMLKGLQHNIVRFYDSWESTYKGCIVLV 300
QY 301 TELMTSGTLTKYLRFRFKMKIKVLRSWCROQLKGLQFLHTPTPIIHRDLKCDNIFFTGP 360
DB 301 TELMTSGTLTKYLRFRFKMKIKVLRSWCROQLKGLQFLHTPTPIIHRDLKCDNIFFTGP 360
QY 361 TGSYKIDGLGATIKRASPAKSVIGTEPFMAPEMYEEXYDSVDVYAFGCMEMATSEY 420
DB 361 TGSYKIDGLGATIKRASPAKSVIGTEPFMAPEMYEEXYDSVDVYAFGCMEMATSEY 420
QY 421 PYSCQNAOITRRVTSGVKSPASFDKVAIPEVKEIIEGCIRONDREYSIKDLNHFQ 480
DB 421 PYSCQNAOITRRVTSGVKSPASFDKVAIPEVKEIIEGCIRONDREYSIKDLNHFQ 480
QY 481 EETGVRELAEEDGEEKIAIKMLRIDIKKLKGYKONDAIEFSPLEBDVPEDAQEM 540
DB 481 EETGVRELAEEDGEEKIAIKMLRIDIKKLKGYKONDAIEFSPLEBDVPEDAQEM 540
QY 541 VESGYVEEGHKTMAKAIKORVSLIKRERQROLVREOEKKKQEBSSLKQVEQSSASQ 600
DB 541 VESGYVEEGHKTMAKAIKORVSLIKRERQROLVREOEKKKQEBSSLKQVEQSSASQ 600
QY 601 TGJKOLPSASTGTPTASTSASVSTOVEPEPEADQOQOYQOPSTSVLSDGTVDSGQ 660
DB 601 TGJKOLPSASTGTPTASTSASVSTOVEPEPEADQOQOYQOPSTSVLSDGTVDSGQ 660
QY 661 SSVETESRVSQOQVSYGSOHEQNAHSTGTPGHIPTVQAQSOQPHGVPPSSVAQSGSQ 720
DB 661 SSVETESRVSQOQVSYGSOHEQNAHSTGTPGHIPTVQAQSOQPHGVPPSSVAQSGSQ 720
QY 721 QPSSSILTVSSSQPIQHPQOQOGIOQTAPPOQTVQYSLSTSTSPATTAQVVSQOPAR 780
DB 721 QPSSSILTVSSSQPIQHPQOQOGIOQTAPPOQTVQYSLSTSTSPATTAQVVSQOPAR 780
QY 781 QVLPVVSAGKQ----- 791
DB 781 QVLPVVSAGKQ----- 791
QY 792 ----- 791
DB 792 ----- 791
QY 841 VSGIPISTPHVSTAGTGFSSLPITMAAGITQPLLTILASSATTAIPGVSTVVSQPLTL 900
DB 841 VSGIPISTPHVSTAGTGFSSLPITMAAGITQPLLTILASSATTAIPGVSTVVSQPLTL 900
QY 901 QPVTOLPSQVHPQLQPAVQSMGI PANTGQAAEVPLSSGDVLYQGFPPRLPQYPGDSNI 960
DB 901 QPVTOLPSQVHPQLQPAVQSMGI PANTGQAAEVPLSSGDVLYQGFPPRLPQYPGDSNI 960
QY 792 ----- 791
DB 792 ----- 791

DB 961 APPSNVAVSVCIHSTVLSPEPMETEVLATPEPTVQVPEVSNLIVPMGVGVQVQVQPG 1020
QY 792 -----STQVSOVAPAEPAVAVQPOQTOTPTTLASSVDSASHDVAAGMS 834
DB 1021 GSIAQAPTTSSQOAVLBSSTQGVSOVAPAEPAVAVQPOQTOTPTTLASSVDSASHDVAAGMS 1080
QY 835 DGENEVSSSGRHHGRRTTKRHYRKSVRSRSHKETSPEKLTILVNSKGDVREVCQLETH 894
DB 1081 DGENEVSSSGRHHGRRTTKRHYRKSVRSRSHKETSPEKLTILVNSKGDVREVCQLETH 1140
QY 895 NKKVATPEKIDGQNPBEIATIMVNDPILAIERESFVDQVREIIEKADMLSEDVSEY 954
DB 1141 NKKVATPEKIDGQNPBEIATIMVNDPILAIERESFVDQVREIIEKADMLSEDVSEY 1200
QY 955 EGDGGLSLQCKDDYGGSGSKLEGEKOPIPASSMQOIGIPSSILTOVYHSGRFTV 1014
DB 1201 EGDGGLSLQCKDDYGGSGSKLEGEKOPIPASSMQOIGIPSSILTOVYHSGRFTV 1260
QY 1015 SPVESRLRESKVPESEITDVAASTQSPQMNLSHASSLSLQAFSELRRAQMTGPN 1074
DB 1261 SPVESRLRESKVPESEITDVAASTQSPQMNLSHASSLSLQAFSELRRAQMTGPN 1320
QY 1075 TAPNFSHTGPTPEVVPPLSSIAGVPTTAATAAPVATSSPNDISTSVYQSEVTPTE 1134
DB 1321 TAPNFSHTGPTPEVVPPLSSIAGVPTTAATAAPVATSSPNDISTSVYQSEVTPTE 1380
QY 1135 EGIAGVASTGVTSGLPIPVSESPVLSVSVSITIPAVSISTSPISLQVPTSEI 1194
DB 1381 EGIAGVASTGVTSGLPIPVSESPVLSVSVSITIPAVSISTSPISLQVPTSEI 1440
QY 1195 VVSTALPSPVTSATSASAGSTATPGPPPAVVSQOAGSTTVAGATILSVSTTSFPS 1224
DB 1441 VVSTALPSPVTSATSASAGSTATPGPPPAVVSQOAGSTTVAGATILSVSTTSFPS 1500
QY 1255 TASQLSLQSSSTPTPLAETVVVSAHSADKTSHTTGLAFSLASBSSSSPAGVSY 1314
DB 1501 TASQLSLQSSSTPTPLAETVVVSAHSADKTSHTTGLAFSLASBSSSSPAGVSY 1560
QY 1315 ISOPGGLHPLVPSVASTPILPQAGPTSTPLLPQVPSIPPLVQPAVAVQOQLIHS 1374
DB 1561 ISOPGGLHPLVPSVASTPILPQAGPTSTPLLPQVPSIPPLVQPAVAVQOQLIHS 1620
QY 1375 QPQPALIPNQHCHCEVSDTOPKAFGIDIKLEBKARSLSSEHSSSGAQAHSVLET 1434
DB 1621 QPQPALIPNQHCHCEVSDTOPKAFGIDIKLEBKARSLSSEHSSSGAQAHSVLET 1680
QY 1435 SLVIBESTVTPGIPITVAAPSKLITSTTCLPPTNLPLGTVALPVTVPVTPQGVSTPVST 1494
DB 1681 SLVIBESTVTPGIPITVAAPSKLITSTTCLPPTNLPLGTVALPVTVPVTPQGVSTPVST 1740
QY 1495 TTSQVAPGTAPSKPEPLTKAPVLPVGTLEPAGTLPSPQLPFPBPSTLQSQOPLIEDDAOL 1554
DB 1741 TTSQVAPGTAPSKPEPLTKAPVLPVGTLEPAGTLPSPQLPFPBPSTLQSQOPLIEDDAOL 1800
QY 1555 RRTLSPEKITTYSAVGPVMAAPTAITTEAGTOPQOKVSVQKEGPVATSSGAGVFXMGFR 1614
DB 1801 RRTLSPEKITTYSAVGPVMAAPTAITTEAGTOPQOKVSVQKEGPVATSSGAGVFXMGFR 1860
QY 1615 QVSVADGAKQEKAKNSBDKSVHPESSSTSESSVSSSPSTLVPEPENGITIPGISD 1674
DB 1861 QVSVADGAKQEKAKNSBDKSVHPESSSTSESSVSSSPSTLVPEPENGITIPGISD 1920
QY 1675 VPESAHTTASEKSDTGOPTKGRQVTTANKVRFVSUKTEKIDITTKKKGPAASP 1734
DB 1921 VPESAHTTASEKSDTGOPTKGRQVTTANKVRFVSUKTEKIDITTKKKGPAASP 1980
QY 1735 FMDLEQAVLPAYIPKKEPELSEPSHLNGPSSDPEAFLSRVDVDSGSPHSPHOLSSGS 1794
DB 1981 FMDLEQAVLPAYIPKKEPELSEPSHLNGPSSDPEAFLSRVDVDSGSPHSPHOLSSGS 2040
QY 1795 LPSQNLSSLSNSFNSSYMSDNESEDIEDDLKELIRLRIDGHLKEIQLOLSRQKHEIES 1854
DB 2041 LPSQNLSSLSNSFNSSYMSDNESEDIEDDLKELIRLRIDGHLKEIQLOLSRQKHEIES 2100

QY 1855 LYTLLGVPPAVIIPPAAPISGRRRRPTKSKGSSSSSSISGNKSPOLSGNLSGQSAASV 1914
DB 2101 LYTLLGVPPAVIIPPAAPISGRRRRPTKSKGSSSSSSISGNKSPOLSGNLSGQSAASV 2160
QY 1915 LHPQOTLHPGNIIPESGONOLLOPLKPSPSNDLYSAFTSDGASIVSLAPGGGTSSTN 1974
DB 2161 LHPQOTLHPGNIIPESGONOLLOPLKPSPSNDLYSAFTSDGASIVSLAPGGGTSSTN 2220
QY 1975 TVGATVNSQAAQAQPPAMTSSRKGTFTDDLHKLVNDNARDAMNISGRGSKGMNYEPG 2034
DB 2221 TVGATVNSQAAQAQPPAMTSSRKGTFTDDLHKLVNDNARDAMNISGRGSKGMNYEPG 2280
QY 2035 MARFSAAPGOLCISMTSNLGGASAPISASATSLGHFTKSMCPPOQYGFPAATPGAKMSGT 2094
DB 2281 MARFSAAPGOLCISMTSNLGGASAPISASATSLGHFTKSMCPPOQYGFPAATPGAKMSGT 2340
QY 2095 GGPAPQPLGQFQPVGTASLQNFNISNLQKSIISNPGSNLRTT 2136
DB 2341 GGPAPQPLGQFQPVGTASLQNFNISNLQKSIISNPGSNLRTT 2382

RESULT 4
US-10-196-935A-2
Sequence 2, Application US/10196935A
Publication No. US20030082720A1
GENERAL INFORMATION:
APPLICANT: Lifton, Richard P
APPLICANT: Wilson, Frederick H
APPLICANT: Choate, Keith
APPLICANT: Ishikawa, Kazuhiko
APPLICANT: Nelson-Williams, Carole
TITLE OF INVENTION: COMPOSITIONS METHODS AND KITS RELATING TO TREATING AND DIAGNOSING
FILE OF INVENTION: HYPERENSION
FILE REFERENCE: 044574-5113
CURRENT APPLICATION NUMBER: US/10/196,935A
CURRENT FILING DATE: 2002-10-25
PRIORITY APPLICATION NUMBER: US 60/306,084
PRIORITY FILING DATE: 2001-07-17
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 2382
TYPE: PRT
ORGANISM: Homo sapiens
US-10-196-935A-2

Query Match 98.8%; Score 10677; DB 14; Length 2382;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;

QY 1 MSGAAEKOSSTPGSLFLSPAPAPKNGSSSDSYGEKLGAAADAVTGRTEETRRRHT 60
DB 1 MSGAAEKOSSTPGSLFLSPAPAPKNGSSSDSYGEKLGAAADAVTGRTEETRRRHT 60
QY 61 MDKSRGAAATTTTTEHFRFRSVICDSNATALELPGPLPSLPOPSIIPAAPQASAPRPH 120
DB 61 MDKSRGAAATTTTTEHFRFRSVICDSNATALELPGPLPSLPOPSIIPAAPQASAPRPH 120
QY 121 REETVATATATSOVAQPPAAAAAGFOAVAPAPSTVPSSTSKDRPVSPSLVSGKEEPP 180
DB 121 REETVATATATSOVAQPPAAAAAGFOAVAPAPSTVPSSTSKDRPVSPSLVSGKEEPP 180
QY 181 ARSGSGGSAKEPQERERQODDIELETKAVGMSNDGRFLKFDIEGRGSKTYKGLD 240
DB 181 ARSGSGGSAKEPQERERQODDIELETKAVGMSNDGRFLKFDIEGRGSKTYKGLD 240
QY 241 TETTVAVMCELQORLTKSRORFKERAEMLKGIQHNIVRFYNSWSTVKGKKCTLV 300
DB 241 TETTVAVMCELQORLTKSRORFKERAEMLKGIQHNIVRFYNSWSTVKGKKCTLV 300
QY 301 TELMTSGTLKTYLKRFRKMKIKVLASMCROILKGLQFLHTRTPPIIHRDLKCDNIFITGP 360
DB 301 TELMTSGTLKTYLKRFRKMKIKVLASMCROILKGLQFLHTRTPPIIHRDLKCDNIFITGP 360

DB 301 TELMTSGTLKTYLKRFRKMKIKVLASMCROILKGLQFLHTRTPPIIHRDLKCDNIFITGP 360
QY 361 TGSVTKIGDLGLATLTKASFAKSVITGPEFMAPEMEKEKYBESVDVYAFGCMEMATSEY 420
DB 361 TGSVTKIGDLGLATLTKASFAKSVITGPEFMAPEMEKEKYBESVDVYAFGCMEMATSEY 420
QY 421 PYSECQMAQIYRRTSGVAPASFDKVALPEVEKEIEGCIKONKDERYSIKDLINHAFO 480
DB 421 PYSECQMAQIYRRTSGVAPASFDKVALPEVEKEIEGCIKONKDERYSIKDLINHAFO 480
QY 481 EETGVARELAEDDGEKIAIKMLRIEDIKLKGKYDNALIESPDLERDVEDVAQEM 540
DB 481 EETGVARELAEDDGEKIAIKMLRIEDIKLKGKYDNALIESPDLERDVEDVAQEM 540
QY 541 VESGYVEGDKHTMAKIKRVSILIKRROUVEEOEKKEESSLQOQVESASQ 600
DB 541 VESGYVEGDKHTMAKIKRVSILIKRROUVEEOEKKEESSLQOQVESASQ 600
QY 601 TGIKOLPSASTGIPASTSTASVSTQVEPEPEADQHQLOQOQPSISVLSGTVDSGQG 660
DB 601 TGIKOLPSASTGIPASTSTASVSTQVEPEPEADQHQLOQOQPSISVLSGTVDSGQG 660
QY 661 SSVFTESRVSQOTVSYGSOHEQAHSTGYVGHIPSTVQAOQOPHYPPSSVAQGSQG 720
DB 661 SSVFTESRVSQOTVSYGSOHEQAHSTGYVGHIPSTVQAOQOPHYPPSSVAQGSQG 720
QY 721 QPSSSLTGVSSQPTOHPOQOQOIQOTAPPOQVOYSLQSTSTSEATTAQPSOPAP 780
DB 721 QPSSSLTGVSSQPTOHPOQOQOIQOTAPPOQVOYSLQSTSTSEATTAQPSOPAP 780
QY 781 QVLPQVSAQKQ----- 791
DB 781 QVLPQVSAQKQPLVSPQVPTIQERPOIPVATQPSVVEVHSGAHLPLPGPLPTLLPOYP 840
QY 792 ----- 791
DB 792 ----- 791
QY 841 VSGIPSTPHVSTAQTFSSLPITMAAGITQPLTLTASATTAIPGVSTVPSQLPTLL 900
DB 841 VSGIPSTPHVSTAQTFSSLPITMAAGITQPLTLTASATTAIPGVSTVPSQLPTLL 900
QY 792 ----- 791
DB 792 ----- 791
QY 901 QVTLQPSQVHPQLLOPAVOSMGIPANLGOAAEVLSSGDLVYGFPRLPQYFGDSNI 960
DB 901 QVTLQPSQVHPQLLOPAVOSMGIPANLGOAAEVLSSGDLVYGFPRLPQYFGDSNI 960
QY 792 ----- 791
DB 792 ----- 791
QY 961 APSNVASVCIHSTVLSBPMPTEVLATPGPEPTVQVYVESNLVPMGVGQVQVSGPQ 1020
DB 961 APSNVASVCIHSTVLSBPMPTEVLATPGPEPTVQVYVESNLVPMGVGQVQVSGPQ 1020
QY 792 -----STGVSOVAPAEPPVAVAOQATQPTTLTASVDSASHVYASMS 834
DB 792 -----STGVSOVAPAEPPVAVAOQATQPTTLTASVDSASHVYASMS 834
QY 1021 GSLAQPSTTSQOAVLSTQVSGVAPAEPPVAVAOQATQPTTLTASVDSASHVYASMS 1080
DB 1021 GSLAQPSTTSQOAVLSTQVSGVAPAEPPVAVAOQATQPTTLTASVDSASHVYASMS 1080
QY 835 DGNENVPSSSGRHHGRTTKHRYKSVSRSHKETSHPKILNVNSKGDVVECOLETH 894
DB 835 DGNENVPSSSGRHHGRTTKHRYKSVSRSHKETSHPKILNVNSKGDVVECOLETH 894
QY 1081 DGNENVPSSSGRHHGRTTKHRYKSVSRSHKETSHPKILNVNSKGDVVECOLETH 1140
DB 1081 DGNENVPSSSGRHHGRTTKHRYKSVSRSHKETSHPKILNVNSKGDVVECOLETH 1140
QY 895 NRKAVTRFKFDLDGNPEIATIMVNDFTLAIRESFVDQVREIIEKADENLSEDSVEP 954
DB 895 NRKAVTRFKFDLDGNPEIATIMVNDFTLAIRESFVDQVREIIEKADENLSEDSVEP 954
QY 1141 NRKAVTRFKFDLDGNPEIATIMVNDFTLAIRESFVDQVREIIEKADENLSEDSVEP 1200
DB 1141 NRKAVTRFKFDLDGNPEIATIMVNDFTLAIRESFVDQVREIIEKADENLSEDSVEP 1200
QY 955 EGDGGLSLQKDYGRSGSKLEGEFRKOPIPASSMPOQIIPFSSLSLQVHVSAGRRTIV 1014
DB 955 EGDGGLSLQKDYGRSGSKLEGEFRKOPIPASSMPOQIIPFSSLSLQVHVSAGRRTIV 1014
QY 1201 EGDGGLSLQKDYGRSGSKLEGEFRKOPIPASSMPOQIIPFSSLSLQVHVSAGRRTIV 1260
DB 1201 EGDGGLSLQKDYGRSGSKLEGEFRKOPIPASSMPOQIIPFSSLSLQVHVSAGRRTIV 1260
QY 1015 SPVPSRRLRESKVPPEITDVAASTAOSPAMNLSHASSLSLQOAFSELRAAQTGEPN 1074
DB 1015 SPVPSRRLRESKVPPEITDVAASTAOSPAMNLSHASSLSLQOAFSELRAAQTGEPN 1074
QY 1261 SPVPSRRLRESKVPPEITDVAASTAOSPAMNLSHASSLSLQOAFSELRAAQTGEPN 1320
DB 1261 SPVPSRRLRESKVPPEITDVAASTAOSPAMNLSHASSLSLQOAFSELRAAQTGEPN 1320
QY 1075 TAPNFSHTGPTPEVVPFSLISAGVPTTAATAAPVPAATSSPNDISTSVIOSEVYVTE 1134
DB 1075 TAPNFSHTGPTPEVVPFSLISAGVPTTAATAAPVPAATSSPNDISTSVIOSEVYVTE 1134
QY 1321 TAPNFSHTGPTPEVVPFSLISAGVPTTAATAAPVPAATSSPNDISTSVIOSEVYVTE 1380
DB 1321 TAPNFSHTGPTPEVVPFSLISAGVPTTAATAAPVPAATSSPNDISTSVIOSEVYVTE 1380
QY 1135 EGIAGVATSTGVVTSGLPIPPVBSBVLSSVSVSITIPAVVVISITTSIPSLQVPTSTSEI 1194
DB 1135 EGIAGVATSTGVVTSGLPIPPVBSBVLSSVSVSITIPAVVVISITTSIPSLQVPTSTSEI 1194
QY 1381 EGIAGVATSTGVVTSGLPIPPVBSBVLSSVSVSITIPAVVVISITTSIPSLQVPTSTSEI 1440
DB 1381 EGIAGVATSTGVVTSGLPIPPVBSBVLSSVSVSITIPAVVVISITTSIPSLQVPTSTSEI 1440

```

QY 1195 VVSSTALPVSUTVSATASAGGSTATPGKPPAVVSQQAAGSTTGATLTSVSTTSPS 1254
Db 1441 VVSSTALPVSUTVSATASAGGSTATPGKPPAVVSQQAAGSTTGATLTSVSTTSPS 1500
QY 1255 TASQSLIOLSSSTSTPLAETVVAHSLDKTSHSTTGLATSLAPSSSSPGAGVSSY 1314
Db 1501 TASQSLIOLSSSTSTPLAETVVAHSLDKTSHSTTGLATSLAPSSSSPGAGVSSY 1560
QY 1315 ISQPGHPLVPSVIASTPILPQAAGPTSTPLPQVPSIPPLVQPVANVPVQOTLHS 1374
Db 1561 ISQPGHPLVPSVIASTPILPQAAGPTSTPLPQVPSIPPLVQPVANVPVQOTLHS 1620
QY 1375 QPQPALPNOPTHCPVDSDTOPKAPGIDIDIKLEEKRLSIFSEHSSGAQHASVLET 1434
Db 1621 QPQPALPNOPTHCPVDSDTOPKAPGIDIDIKLEEKRLSIFSEHSSGAQHASVLET 1680
QY 1435 SLVSESTVPGIPTTAVAAPSKLLTSTSTCLPPTNLPGTVALPPTPVVTGQVSTPVST 1494
Db 1681 SLVSESTVPGIPTTAVAAPSKLLTSTSTCLPPTNLPGTVALPPTPVVTGQVSTPVST 1740
QY 1495 TTSGVKGPTAPSKPELTAPVLPVGTLPAGTLPSBOLPPPGPSLTQSQOPLBDDAOL 1554
Db 1741 TTSGVKGPTAPSKPELTAPVLPVGTLPAGTLPSBOLPPPGPSLTQSQOPLBDDAOL 1800
QY 1555 RRTLSPEKITVTSVAVGVMAAPTAITEAGTQPOKGVSVQKEGVLATSSGAGVFKMGF 1614
Db 1801 RRTLSPEKITVTSVAVGVMAAPTAITEAGTQPOKGVSVQKEGVLATSSGAGVFKMGF 1860
QY 1615 QVSVAAQDAQKEGNKSDAKSVHSESTSSSVLSSSPSTLVKPPNITTPGLSSD 1674
Db 1861 QVSVAAQDAQKEGNKSDAKSVHSESTSSSVLSSSPSTLVKPPNITTPGLSSD 1920
QY 1675 VPESAHKTTASEAASDITGQPTKVGRCVTTTANKVGRSVKTEDKIDTCKECPVASSP 1734
Db 1921 VPESAHKTTASEAASDITGQPTKVGRCVTTTANKVGRSVKTEDKIDTCKECPVASSP 1980
QY 1735 FMDLEQAVLPVAVIPEKKEPELSEPSHLNGPSSDPEAAFLSRDVGSGSPHSPIQLSSKS 1794
Db 1981 FMDLEQAVLPVAVIPEKKEPELSEPSHLNGPSSDPEAAFLSRDVGSGSPHSPIQLSSKS 2040
QY 1795 LPSQNLGOSISNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 1854
Db 2041 LPSQNLGOSISNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNSNS 2100
QY 1855 LYTLLGKVPVAVIIPPAAPLSGRRRRPTKSKGSKSSRSSSLGKNSPOLSGNLGQSSASV 1914
Db 2101 LYTLLGKVPVAVIIPPAAPLSGRRRRPTKSKGSKSSRSSSLGKNSPOLSGNLGQSSASV 2160
QY 1915 LHPQOTLHPGNIPESSQNLLOPLKPSRSSDNLYSAFSTSGAISVPLSLAPGQSTSTN 1974
Db 2161 LHPQOTLHPGNIPESSQNLLOPLKPSRSSDNLYSAFSTSGAISVPLSLAPGQSTSTN 2220
QY 1975 TVGATVNSQAQAOPPMTSSRKGTFTDDHLKLVDMNARDAMNLSGRRGSGHNNYGP 2034
Db 2221 TVGATVNSQAQAOPPMTSSRKGTFTDDHLKLVDMNARDAMNLSGRRGSGHNNYGP 2280
QY 2035 MARKFSAPGOLCISMTSNLGGSAFISASAATSLGHFTKSMCPPOQYFPAITPFGAOMSGT 2094
Db 2281 MARKFSAPGOLCISMTSNLGGSAFISASAATSLGHFTKSMCPPOQYFPAITPFGAOMSGT 2340
QY 2095 GGPAPQPLQGFQPVGTASLQNFNISNLQKSIINPGSNLRTT 2136
Db 2341 GGPAPQPLQGFQPVGTASLQNFNISNLQKSIINPGSNLRTT 2382

```

```

? APPLICANT: Casman, Stacie
? APPLICANT: Colman, Steven
? APPLICANT: Edinger, Shlomit R.
? APPLICANT: Ellerman, Karen
? APPLICANT: Gerlach, Valerie
? APPLICANT: Gunther, Erik
? APPLICANT: Kekuda, Ramesh
? APPLICANT: MacDougall, John R.
? APPLICANT: Mehrahan, Fuad
? APPLICANT: Paturajan, Meera
? APPLICANT: Rothenberg, Mark
? APPLICANT: Shinkets, Richard
? APPLICANT: Smithson, Glennda
? APPLICANT: Spytek, Kimberly A.
? APPLICANT: Stone, David J.
? APPLICANT: Verneet, Corine A.M.
? APPLICANT: Zethusen, Bryan D.
? TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
? FILE REFERENCE: 21402-250 (CURA-550)
? CURRENT FILING DATE: 2002-12-09
? PRIOR APPLICATION NUMBER: 60/262,454
? PRIOR FILING DATE: 2001-01-18
? PRIOR APPLICATION NUMBER: 60/272,920
? PRIOR FILING DATE: 2001-03-02
? PRIOR APPLICATION NUMBER: 60/284,549
? PRIOR FILING DATE: 2001-04-18
? PRIOR APPLICATION NUMBER: 60/303,229
? PRIOR FILING DATE: 2001-07-05
? PRIOR APPLICATION NUMBER: 60/262,892
? PRIOR FILING DATE: 2001-01-19
? PRIOR APPLICATION NUMBER: 60/263,605
? PRIOR FILING DATE: 2001-01-23
? PRIOR APPLICATION NUMBER: 60/269,098
? PRIOR FILING DATE: 2001-02-15
? PRIOR APPLICATION NUMBER: 60/264,159
? PRIOR FILING DATE: 2001-01-25
? PRIOR APPLICATION NUMBER: 60/265,517
? PRIOR FILING DATE: 2001-01-31
? PRIOR APPLICATION NUMBER: 60/271,855
? PRIOR FILING DATE: 2001-02-27
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 97
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO: 40
? LENGTH: 2382
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-052-648A-40
Query Match 98.8%; Score 10677; DB 15; Length 2382;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;
QY 1 MSGGAEEKOSTGSLPLSPAPAPKXGSSSDSVGEKLGAAADAVGTETERRRRRT 60
Db 1 MSGGAEEKOSTGSLPLSPAPAPKXGSSSDSVGEKLGAAADAVGTETERRRRRT 60
QY 61 MDKDSRGAATTTTTEHRRFRRSVTCDSNATALELPGLPLSLPQPSIPAAVPOSAPPEPH 120
Db 61 MDKDSRGAATTTTTEHRRFRRSVTCDSNATALELPGLPLSLPQPSIPAAVPOSAPPEPH 120
QY 121 REETVATATNSQVAAQOPPAAGAAQVAGPASTVSSSKRPVSQPSLVGSKKEPP 180
Db 121 REETVATATNSQVAAQOPPAAGAAQVAGPASTVSSSKRPVSQPSLVGSKKEPP 180
QY 181 ARSGSGGSAKEQOEERSQOQDDIELEETAQVAGSNDGRFLKPDIEIGRSFKTVVYGLD 240
Db 181 ARSGSGGSAKEQOEERSQOQDDIELEETAQVAGSNDGRFLKPDIEIGRSFKTVVYGLD 240
QY 241 TETTVAVAGCELODRKLYTSERQFKEEAEMLXGLQHPNIVRFYDSWESTVKGKCTIVV 300

```

```

RESULT 5
US-10-052-648A-40
; Sequence 40, Application US/10052648A
; Publication No. US2004000558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine

```

```

Db 241 TETTVAVAMELODRKLTKEBROFKEBAEMLKGLQHNTIVRFYDSESTYVKGKCCIVLV 300
QY 301 TELMTSGTLKTYLKRFRVMKIKVLRSWCROJLKLQPLHRTPTPIHRDLKCNIPITGP 360
Db 301 TELMTSGTLKTYLKRFRVMKIKVLRSWCROJLKLQPLHRTPTPIHRDLKCNIPITGP 360
QY 361 TGSYKIDIGLGLATLKRASPAKSVIGTPEFMAPEMYEKEKDSVVVYAFGCMLEMAISEY 420
Db 361 TGSYKIDIGLGLATLKRASPAKSVIGTPEFMAPEMYEKEKDSVVVYAFGCMLEMAISEY 420
QY 421 PYSCQNAQAQIYRRVTSVSGVSPASFDKVAIPVKEIIEBCIRQNDERYSIKDLNHAFFQ 480
Db 421 PYSCQNAQAQIYRRVTSVSGVSPASFDKVAIPVKEIIEBCIRQNDERYSIKDLNHAFFQ 480
QY 481 EETGVAVELAEEDDGEKIAIKLMTIRIEDIKLKGYKXONEAIEFSFDLERVPEVDAQEM 540
Db 481 EETGVAVELAEEDDGEKIAIKLMTIRIEDIKLKGYKXONEAIEFSFDLERVPEVDAQEM 540
QY 541 VESGYVEGDHKTMAKAIKDRVSLIKRREOROLVREBQEKKOESLSLKQOVQSSASQ 600
Db 541 VESGYVEGDHKTMAKAIKDRVSLIKRREOROLVREBQEKKOESLSLKQOVQSSASQ 600
QY 601 TGICOLPASASTGITPTASTTSASVSTOVEPEBEPADQHOLQYQOPSISVLSGTVDSGQG 660
Db 601 TGICOLPASASTGITPTASTTSASVSTOVEPEBEPADQHOLQYQOPSISVLSGTVDSGQG 660
QY 661 SSVTEBESVSSQOQVSYSGHBOAHSTGTVPCHIPSTYQAOSONHGVYPPSSVAAQOGSOG 720
Db 661 SSVTEBESVSSQOQVSYSGHBOAHSTGTVPCHIPSTYQAOSONHGVYPPSSVAAQOGSOG 720
QY 721 QPSSSLSLTVGSSSQPIQHPOQOQGIQCTAPPOQTQVYSLSQTSSEATTTQAPVSPQAP 780
Db 721 QPSSSLSLTVGSSSQPIQHPOQOQGIQCTAPPOQTQVYSLSQTSSEATTTQAPVSPQAP 780
QY 781 QVLPQVSAKXQ----- 791
Db 781 QVLPQVSAKXQ----- 791
QY 791 ----- 791
Db 791 ----- 791
QY 841 VSOIPISTPHVSTAQTGFSSLPITMAAGITOPLLTLASSATTAAIPGVSTVPSQLPTLL 900
Db 841 VSOIPISTPHVSTAQTGFSSLPITMAAGITOPLLTLASSATTAAIPGVSTVPSQLPTLL 900
QY 792 ----- 791
Db 792 ----- 791
QY 901 QPVTQVPSQVHPOLLQPAVQSMGIRANLQQAHEVLSGDVLYGCFPRRLPQVPGDSNI 960
Db 901 QPVTQVPSQVHPOLLQPAVQSMGIRANLQQAHEVLSGDVLYGCFPRRLPQVPGDSNI 960
QY 792 ----- 791
Db 792 ----- 791
QY 961 APSSNVASVCIHSTVYLSPPMPTVLA TPGYFTVVQPYVESNLVPMGVGQGVQVQSOG 1020
Db 961 APSSNVASVCIHSTVYLSPPMPTVLA TPGYFTVVQPYVESNLVPMGVGQGVQVQSOG 1020
QY 792 ----- STGVSYVAPABEVAVAOPOATOPTTLASSYDASHDYASMS 834
Db 792 ----- STGVSYVAPABEVAVAOPOATOPTTLASSYDASHDYASMS 834
QY 1021 GSLAQAPTTSSQOAVLESTQGVSAPEVAVAOPOATOPTTLASSYDASHDYASMS 1080
Db 1021 GSLAQAPTTSSQOAVLESTQGVSAPEVAVAOPOATOPTTLASSYDASHDYASMS 1080
QY 835 DGENEVPSSSGRHEGRATTKRHYRKSVRSRSHKESRPEKRLILNVSNKGRDVECOLETH 894
Db 835 DGENEVPSSSGRHEGRATTKRHYRKSVRSRSHKESRPEKRLILNVSNKGRDVECOLETH 894
QY 1081 DGENEVPSSSGRHEGRATTKRHYRKSVRSRSHKESRPEKRLILNVSNKGRDVECOLETH 1140
Db 1081 DGENEVPSSSGRHEGRATTKRHYRKSVRSRSHKESRPEKRLILNVSNKGRDVECOLETH 1140
QY 895 NRKAVTFKFDLDGNDPEEIIATMVNDPILAIERBSFYDVOREIIEKADEMISEDVSEP 954
Db 895 NRKAVTFKFDLDGNDPEEIIATMVNDPILAIERBSFYDVOREIIEKADEMISEDVSEP 954
QY 1141 NRKAVTFKFDLDGNDPEEIIATMVNDPILAIERBSFYDVOREIIEKADEMISEDVSEP 1200
Db 1141 NRKAVTFKFDLDGNDPEEIIATMVNDPILAIERBSFYDVOREIIEKADEMISEDVSEP 1200
QY 955 EGDQGLESLQGDQVGFSGSQKLBEFQKOPTPASMPOQIGITPSLQOVHSAQRRTIV 1014
Db 955 EGDQGLESLQGDQVGFSGSQKLBEFQKOPTPASMPOQIGITPSLQOVHSAQRRTIV 1014
QY 1201 EGDQGLESLQGDQVGFSGSQKLBEFQKOPTPASMPOQIGITPSLQOVHSAQRRTIV 1260
Db 1201 EGDQGLESLQGDQVGFSGSQKLBEFQKOPTPASMPOQIGITPSLQOVHSAQRRTIV 1260
QY 1015 SPVESRRLRESKVFSEITDTVAASTAOSPGMNLSSHASLSLQOAFSELRRQAQMTBEPN 1074
Db 1015 SPVESRRLRESKVFSEITDTVAASTAOSPGMNLSSHASLSLQOAFSELRRQAQMTBEPN 1074
QY 1261 SPVESRRLRESKVFSEITDTVAASTAOSPGMNLSSHASLSLQOAFSELRRQAQMTBEPN 1320
Db 1261 SPVESRRLRESKVFSEITDTVAASTAOSPGMNLSSHASLSLQOAFSELRRQAQMTBEPN 1320
QY 1075 TAPPNFHTGPTFPVPPFLSSIAQVPTTAATAATAPVATSSPNDISTSVIQSEVTPTE 1134
Db 1075 TAPPNFHTGPTFPVPPFLSSIAQVPTTAATAATAPVATSSPNDISTSVIQSEVTPTE 1134
QY 1321 TAPPNFHTGPTFPVPPFLSSIAQVPTTAATAATAPVATSSPNDISTSVIQSEVTPTE 1380
Db 1321 TAPPNFHTGPTFPVPPFLSSIAQVPTTAATAATAPVATSSPNDISTSVIQSEVTPTE 1380

```

```

QY 1135 EGIAGVATSGVTSGLPIPVSESPVLSVVSITITPAVVSISTSPSLQVPTSEI 1194
Db 1381 EGIAGVATSGVTSGLPIPVSESPVLSVVSITITPAVVSISTSPSLQVPTSEI 1440
QY 1195 VVSSTALYPVTVSATSASAGSSTATPGPKPVPVVSQQAAGSTVGAITLVSSTTSFPS 1254
Db 1441 VVSSTALYPVTVSATSASAGSSTATPGPKPVPVVSQQAAGSTVGAITLVSSTTSFPS 1500
QY 1255 TASQOLSTLSSSTPTPLAETTVVSANSLDKTSHSSTTGALFSLSASSSSSPQAGVSY 1314
Db 1501 TASQOLSTLSSSTPTPLAETTVVSANSLDKTSHSSTTGALFSLSASSSSSPQAGVSY 1560
QY 1315 ISOPGHLPLVTVSVIATSTPLPQAPGPTSTPLLPQVPSIPELVQPVANVAVOQTLHS 1374
Db 1561 ISOPGHLPLVTVSVIATSTPLPQAPGPTSTPLLPQVPSIPELVQPVANVAVOQTLHS 1620
QY 1375 QPOPALLPNQPHTHCEVSDTOPKABQIDIDIKLLEKRLSFESEHSSGAQHASVSIET 1434
Db 1621 QPOPALLPNQPHTHCEVSDTOPKABQIDIDIKLLEKRLSFESEHSSGAQHASVSIET 1680
QY 1435 SLVIESVTVTGIPPTTAVAPSKULTSTSTCLPPTNLPLGTVALVTVVTPVPGVSTPST 1494
Db 1681 SLVIESVTVTGIPPTTAVAPSKULTSTSTCLPPTNLPLGTVALVTVVTPVPGVSTPST 1740
QY 1495 TTSGVKPGTAPSKPPLTKAPVLPVGTLPAGTLPSBOLPPPGBSLTQSOQPLEDLDAOL 1554
Db 1741 TTSGVKPGTAPSKPPLTKAPVLPVGTLPAGTLPSBOLPPPGBSLTQSOQPLEDLDAOL 1800
QY 1555 RRLTSPKXITVTSVAVGVSNAAPTAITBAGTOPQKGVSYQVKEGVTLATSSGAGVFKMGFR 1614
Db 1801 RRLTSPKXITVTSVAVGVSNAAPTAITBAGTOPQKGVSYQVKEGVTLATSSGAGVFKMGFR 1860
QY 1615 QVSYAADQAOKEGKNKEDAKSVFESESTSESVLSSSSPSTLVKPEPNCITTPGISSD 1674
Db 1861 QVSYAADQAOKEGKNKEDAKSVFESESTSESVLSSSSPSTLVKPEPNCITTPGISSD 1920
QY 1675 VPESAHKTITASAKSDPTGQPTKVGRFQVTTTANKVGRFVSXKTEDKITDTKKEGPVAPSP 1734
Db 1921 VPESAHKTITASAKSDPTGQPTKVGRFQVTTTANKVGRFVSXKTEDKITDTKKEGPVAPSP 1980
QY 1735 FMDLEQAVLPAVDPKKEPELSEPSHNGPSSDPEAFLSHDVDDGSGSPHSPQLSSKS 1794
Db 1981 FMDLEQAVLPAVDPKKEPELSEPSHNGPSSDPEAFLSHDVDDGSGSPHSPQLSSKS 2040
QY 1795 LPSQNLQSLSNSPNSSYMSSDNESDIEDDLKLELRLRDKHLEIIDLQSRQHEIES 1854
Db 2041 LPSQNLQSLSNSPNSSYMSSDNESDIEDDLKLELRLRDKHLEIIDLQSRQHEIES 2100
QY 1855 LYTILGKVPRAVITPPAAPLISGRRRRPTKSKGSSSSSSSILGNKSPQLSMLSGQSAASV 1914
Db 2101 LYTILGKVPRAVITPPAAPLISGRRRRPTKSKGSSSSSSSILGNKSPQLSMLSGQSAASV 2160
QY 1915 LHPQOHTHPNPNIPESGONOLLOPLKPSBSSDNLYSATFSGAISVPSLSAGQSTSTN 1974
Db 2161 LHPQOHTHPNPNIPESGONOLLOPLKPSBSSDNLYSATFSGAISVPSLSAGQSTSTN 2220
QY 1975 TVGATVNSQAQAOPAMTSSRKGFITDDLKLVDNWRDAMNISGRGSGGHNMYEPPG 2034
Db 2221 TVGATVNSQAQAOPAMTSSRKGFITDDLKLVDNWRDAMNISGRGSGGHNMYEPPG 2280
QY 2035 MARKFSAPGOLCISMTSNLGGSAPIISAASATSLGFTXSMCPQOYGPATPFGAOWSGT 2094
Db 2281 MARKFSAPGOLCISMTSNLGGSAPIISAASATSLGFTXSMCPQOYGPATPFGAOWSGT 2340
QY 2095 GGPAPOPLGOFQPVGTASLQNFNINSLNLOKSIINPPGSMLRTT 2136
Db 2341 GGPAPOPLGOFQPVGTASLQNFNINSLNLOKSIINPPGSMLRTT 2382

```

RESULT 6
 US-10-408-765A-1404
 ; Sequence 1404, Application US/10408765A

Publication No. US20040101674A1
 GENERAL INFORMATION:
 APPLICANT: Ghosh, Soumitra S.
 APPLICANT: Fahy, Boia D.
 APPLICANT: Zhang, Bing
 APPLICANT: Gibson, Bradford W.
 APPLICANT: Taylor, Steven W.
 APPLICANT: Glenn, Gary M.
 APPLICANT: Marnock, Dale E.
 TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 FILE REFERENCE: 660088.465
 CURRENT APPLICATION NUMBER: US/10/408,765A
 CURRENT FILING DATE: 2003-04-04
 NUMBER OF SEQ ID NOS: 3077
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1404
 LENGTH: 2382
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-408-765A-1404

Query Match 98.8%; Score 10677; DB 16; Length 2382;
 Best Local Similarity 89.6%; Pred. No. 0;
 Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;

QY 1 MSGAAEKQKSTPSSLPSPAPAPKNGSSSDSSVGEKLGAAADAVTGRTEYRRRH 60
 Db 1 MSGAAEKQKSTPSSLPSPAPAPKNGSSSDSSVGEKLGAAADAVTGRTEYRRRH 60
 QY 61 MDKSRGAAATTTTTHRRFFRRSVICSNATALELPGIPSLPPSPIDPAVPGAPBPH 120
 Db 61 MDKSRGAAATTTTTHRRFFRRSVICSNATALELPGIPSLPPSPIDPAVPGAPBPH 120
 QY 121 REETVATATSVQVAAQPPAAAPGEQAVAPAPSTVSTSKDMPVQPSIVSGKEPP 180
 Db 121 REETVATATSVQVAAQPPAAAPGEQAVAPAPSTVSTSKDMPVQPSIVSGKEPP 180
 QY 121 REETVATATSVQVAAQPPAAAPGEQAVAPAPSTVSTSKDMPVQPSIVSGKEPP 180
 Db 121 REETVATATSVQVAAQPPAAAPGEQAVAPAPSTVSTSKDMPVQPSIVSGKEPP 180
 QY 181 ARSSGGGSAKEPEBERSQOQDIEELETKAVGMSNDGRFLKPDIEIGRSGFKTVYKGLD 240
 Db 181 ARSSGGGSAKEPEBERSQOQDIEELETKAVGMSNDGRFLKPDIEIGRSGFKTVYKGLD 240
 QY 241 TETVEVAMCELOQRKLTKEBRQFKKEAEMLKGLQPNIVRPFDSNESTYKGGKIVLV 300
 Db 241 TETVEVAMCELOQRKLTKEBRQFKKEAEMLKGLQPNIVRPFDSNESTYKGGKIVLV 300
 QY 241 TETVEVAMCELOQRKLTKEBRQFKKEAEMLKGLQPNIVRPFDSNESTYKGGKIVLV 300
 Db 241 TETVEVAMCELOQRKLTKEBRQFKKEAEMLKGLQPNIVRPFDSNESTYKGGKIVLV 300
 QY 301 TELMTSGTLKTYLKRFFVMKIKVLRSCROIKGLQHLHTRTPPIIHRDLKCONIFITGP 360
 Db 301 TELMTSGTLKTYLKRFFVMKIKVLRSCROIKGLQHLHTRTPPIIHRDLKCONIFITGP 360
 QY 361 TGSVYKIGDLGLATLKRASFAKSVIETPEFMAPEMEYKEDSYDVVAFGCMLEMATSEY 420
 Db 361 TGSVYKIGDLGLATLKRASFAKSVIETPEFMAPEMEYKEDSYDVVAFGCMLEMATSEY 420
 QY 421 PYSECQAAQIVRRVTSVGVKPAFPDVAIPEVKEIIEGCTRQKDEKYSIKDLINAFQ 480
 Db 421 PYSECQAAQIVRRVTSVGVKPAFPDVAIPEVKEIIEGCTRQKDEKYSIKDLINAFQ 480
 QY 481 EETGVEVLAEEEDGKIAIKMLRIEDIKLKGAKYKNDNAIFPSFDEVEDVAQEM 540
 Db 481 EETGVEVLAEEEDGKIAIKMLRIEDIKLKGAKYKNDNAIFPSFDEVEDVAQEM 540
 QY 541 VESGYICEGDHKTMAKAIKDRVSLIKRKEQRLVBEQEKKESSLSKQVEGSSAQ 600
 Db 541 VESGYICEGDHKTMAKAIKDRVSLIKRKEQRLVBEQEKKESSLSKQVEGSSAQ 600
 QY 601 TGIKOLPSASTGTPTASTASVSTOVEBEPEADHQOLQYQPSISVLSDTVDSGG 660
 Db 601 TGIKOLPSASTGTPTASTASVSTOVEBEPEADHQOLQYQPSISVLSDTVDSGG 660
 QY 661 SSVFTESRVSSQQTIVSGSHQHBASTGTVPGHIPSTVQAQSPHGVYPPSSVAQSG 720
 Db 661 SSVFTESRVSSQQTIVSGSHQHBASTGTVPGHIPSTVQAQSPHGVYPPSSVAQSG 720

QY 721 QPSSSLTGVSASSQPIQHPOQOQGIQOTAPPOQTQVSLSQSTSSSEATTAPQVQOPAP 780
 Db 721 QPSSSLTGVSASSQPIQHPOQOQGIQOTAPPOQTQVSLSQSTSSSEATTAPQVQOPAP 780
 QY 781 QVLPQVSAGKQ----- 791
 Db 781 QVLPQVSAGKQLPVSPVPTIQGEPIQVATQPSVVPVHSGAHFLPVGQPLPTPLLPQY 840
 QY 792 ----- 791
 Db 841 VSGPISTPHVSTAQTGSSSLPITMAAGITQPLTLASSATTAAIPGVSTVPQSLPTLL 900
 QY 792 ----- 791
 Db 901 QPVTQLSPQHOLLQPAVQSMGIPANIGQAAEVLSSGDLVYGFPRPLPQYQGSNI 960
 QY 792 ----- 791
 Db 961 APSSNVASVCIHSTVLSPPMPTVLAITPGYPTVQPVESNLLVPMGCVGGQVQVQOPG 1020
 QY 792 -----STQGVQVADPAEPVANAQPOATQPTTLASSVDSASHSDVAGMS 834
 Db 1021 GSIAQAPETSSQAVLSTQGVQVADPAEPVANAQPOATQPTTLASSVDSASHSDVAGMS 1080
 QY 835 DGENEVSSSGRHEGRTTKRHYRKSRSRHEKTSRPKLILNWSKGRVVECOLETH 894
 Db 1081 DGENEVSSSGRHEGRTTKRHYRKSRSRHEKTSRPKLILNWSKGRVVECOLETH 1140
 QY 895 NRKMTPEKFDLDGDNPEIATIMVNDPIAIBESFVDQVREIIEKADWMLSDVSE 954
 Db 1141 NRKMTPEKFDLDGDNPEIATIMVNDPIAIBESFVDQVREIIEKADWMLSDVSE 1200
 QY 955 EGDQGLSLGKXDYGGSSQKLEGEFKQPIPASSMPQOIGIPTSLTQVHVSAGRPTV 1014
 Db 1201 EGDQGLSLGKXDYGGSSQKLEGEFKQPIPASSMPQOIGIPTSLTQVHVSAGRPTV 1260
 QY 1015 SPYBESLRKSKYFPEITDTVAASTQPSGMLSHASSLSIQAASELRRAQMTGPN 1074
 Db 1261 SPYBESLRKSKYFPEITDTVAASTQPSGMLSHASSLSIQAASELRRAQMTGPN 1320
 QY 1075 TAPNFSHTGTPPVVPPFLSSIAGVPTTAATAAPVATSSPNDISTSVIOSEVTPTE 1134
 Db 1321 TAPNFSHTGTPPVVPPFLSSIAGVPTTAATAAPVATSSPNDISTSVIOSEVTPTE 1380
 QY 1135 EGIAGVATSTGVTSGGLPIPVSESPVUSVSSITIPAVASISTTSPSLQVPTSTSEI 1194
 Db 1381 EGIAGVATSTGVTSGGLPIPVSESPVUSVSSITIPAVASISTTSPSLQVPTSTSEI 1440
 QY 1195 VVSTALYPSVTVSATSSAGSTATPCKPAPVAVQOAAAGSTTVGATLTSVSTTSPFS 1254
 Db 1441 VVSTALYPSVTVSATSSAGSTATPCKPAPVAVQOAAAGSTTVGATLTSVSTTSPFS 1500
 QY 1255 TASQLSLOLSSSTPTLAEVTVVASHSLDKTSHSSTGLAFSLAPSSSSSGAGVSSY 1314
 Db 1501 TASQLSLOLSSSTPTLAEVTVVASHSLDKTSHSSTGLAFSLAPSSSSSGAGVSSY 1560
 QY 1315 ISOPGHLVLVPSVASTPILPOAAGPTSTPLLPVPSIPVLQVAVANVAVAOQLTHS 1374
 Db 1561 ISOPGHLVLVPSVASTPILPOAAGPTSTPLLPVPSIPVLQVAVANVAVAOQLTHS 1620
 QY 1375 QPQPALPQPHTHCEVNDTOPKAPGIDDKTEEEKRSLFSEHSSSGAGHASTSLT 1434
 Db 1621 QPQPALPQPHTHCEVNDTOPKAPGIDDKTEEEKRSLFSEHSSSGAGHASTSLT 1680
 QY 1435 SLVIESTVTPGIPPTAVAPSKLLSTSTCLPPTMLPLGTVALPVVPTVPGQVSTPVST 1494
 Db 1681 SLVIESTVTPGIPPTAVAPSKLLSTSTCLPPTMLPLGTVALPVVPTVPGQVSTPVST 1740
 QY 1495 TTSVYKPGTAPSKPLTKAPVLPVGTGLPAGLPSBOLPPFPGPSITQSQOPLDLDLQ 1554
 Db 1741 TTSVYKPGTAPSKPLTKAPVLPVGTGLPAGLPSBOLPPFPGPSITQSQOPLDLDLQ 1800

QY	1555	RRLTSPXKITIVTAVGVPVSMAPALITBAGPOPKGVSYQVKEGVLTSSAGVFKMGRF	1614
Db	1801	RRLTSPKITIVTAVGVPVSMAPALITBAGTOPKGVSYQVKEGVLTSSAGVFKMGRF	1860
QY	1615	QVSAVAADGAQKEGKNKSEDAKSVHFEESTSSSVLSSSSPESTLVKEBPNGITIPGISD	1674
Db	1861	QVSAVAADGAQKEGKNKSEDAKSVHFEESTSSSVLSSSSPESTLVKEBPNGITIPGISD	1920
QY	1675	VPESAHTTASSEAKSDTGOPTKVGRFOVYTTANKVGRFSVKTDKXITDYTKKEGPVSP	1734
Db	1921	VPESAHTTASSEAKSDTGOPTKVGRFOVYTTANKVGRFSVKTDKXITDYTKKEGPVSP	1980
QY	1735	FMDLEQAVLPAVITPKKKEPELSEPSHLNGPSSDPEAALISDVDDGSSPHSPHLSKS	1794
Db	1981	FMDLEQAVLPAVITPKKKEPELSEPSHLNGPSSDPEAALISDVDDGSSPHSPHLSKS	2040
QY	1795	LPSONLQSLSNSFNSSYMSSDNESDIEDDEJLKLRLRDKHLEKEIDOLQSRQKHIES	1854
Db	2041	LPSONLQSLSNSFNSSYMSSDNESDIEDDEJLKLRLRDKHLEKEIDOLQSRQKHIES	2100
QY	1855	LYTKLGKVPVAVITPPAAPLSGRRRRPPTKSGKSYSSRSSLSGKNSPOLSGNLSNGSASV	1914
Db	2101	LYTKLGKVPVAVITPPAAPLSGRRRRPPTKSGKSYSSRSSLSGKNSPOLSGNLSNGSASV	2160
QY	1915	LHPQOULHPPCNIPESGONOLLOPKSPSSDNLYSAFTSPGALISVPELSAPGQSTSTN	1974
Db	2161	LHPQOULHPPCNIPESGONOLLOPKSPSSDNLYSAFTSPGALISVPELSAPGQSTSTN	2222
QY	1975	TVGATVNSQAQAOPPAMTSSSRKGTFTDHLKLVDMNARDAMNLSGRRGSKGHMNYECPG	2034
Db	2221	TVGATVNSQAQAOPPAMTSSSRKGTFTDHLKLVDMNARDAMNLSGRRGSKGHMNYECPG	2288
QY	2035	MARKPSAPGOLCISMTSNLGSAPISAAASATSLGHFTSMCPPOQYGPATPFGQWNGT	2099
Db	2281	MARKPSAPGOLCISMTSNLGSAPISAAASATSLGHFTSMCPPOQYGPATPFGQWNGT	2340
QY	2095	GGPAPQPLGQFOPVGTASLQNFNNISNLKSTSNPBGSLRTT	2136
Db	2341	GGPAPQPLGQFOPVGTASLQNFNNISNLKSTSNPBGSLRTT	2382

```

RESULT 7
US-10-052-648A-39
: Sequence 39, Application US/10052648A
: Publication No. US20040005568A1
: GENERAL INFORMATION:
: APPLICANT: Anderson, David
: APPLICANT: Burgess, Catherine
: APPLICANT: Casman, Stacie
: APPLICANT: Colman, Steven
: APPLICANT: Edinger, Shlomit R.
: APPLICANT: Ellerman, Karen
: APPLICANT: Gerlach, Valerie
: APPLICANT: Gunther, Erik
: APPLICANT: Kekuda, Ramesh
: APPLICANT: MacDougall, John R.
: APPLICANT: Mehrtaban, Fuad
: APPLICANT: Paturajan, Meera
: APPLICANT: Rothenberg, Mark
: APPLICANT: Shinkets, Richard
: APPLICANT: Smithson, Glenda
: APPLICANT: Spytek, Kimberly A.
: APPLICANT: Stone, David J.
: APPLICANT: Vernet, Corine A.M.
: APPLICANT: Zetnusen, Bryan D.
: TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
: TITLE OF INVENTION: USING THE SAME
: FILE REFERENCE: 21402-250 (CURA-550)
: CURRENT APPLICATION NUMBER: US/10/052,648A
: CURRENT FILING DATE: 2002-12-09
: PRIOR APPLICATION NUMBER: 60/262,454
: PRIOR FILING DATE: 2001-01-18
: PRIOR APPLICATION NUMBER: 60/272,920

```

[illegible]

QY 658 GQSSVFTESRVSQOQTVSYSGOHEOAHSTGTVPHHISTVOAQSGPHGVYPPSSVAGQ 717
 DB 659 GQSSVFTESRVSQOQTVSYSGOHEOAHSTGTVPHHISTVOAQSGPHGVYPPSSVAGQ 718
 QY 718 SQGQSSSLTGVSQSOPIQHPOQOQGIQOAPPOQTVYSLQSTSTSEKTTAQVSP 777
 DB 719 MOGQP-SSSLAGVYSSQVQHP-00QGIQTPVPOQAVQVSLPQAASSSEG-TVOQPSQ- 774
 QY 778 QAPVLPQVSAQKQ-STQGVSAQVAPAEVAVAPQATQPTLASSVDSASHDVAGMSDG 836
 DB 775 -----PQVSAQSTQSTQGVSAQAPPEBQTPITQSQPQVPLVSSVDSASHDVAGMSDG 828
 QY 837 NENAPSSSGRHEGTTKRYKSVRSRHEKTSRPKLRILNVSNKGDRVVECOLETHNR 896
 DB 829 NENAPSSSGRHEGTTKRYKSVRSRHEKTSRPKLRILNVSNKGDRVVECOLETHNR 888
 QY 897 KMTVFKFDLQDNDNEELATIMVNDLILATRESFVQVRIIEKABMLSEDEVBE 956
 DB 889 KMTVFKFDLQDNDNEELATIMVNDLILATRESFVQVRIIEKABMLSEDEVBE 948
 QY 957 DQGLESLQKQDYGFSQSGKLEGEFKQPIPASMPQOIGIPTSSLTOVHSAGRFTVSP 1016
 DB 949 DQGLESLQKQDYGFSQSGKLEGEFKQPIPASMPQOIGIPTSSLTOVHSAGRFTVSP 1008
 QY 1017 VPBRLESEKVFPEITDTVAASQSPGMNLSHASSLSLOQAFSELRAQMTGENT 1076
 DB 1009 VPBRLESEKVFPEITDTVAASQSPGMNLSHASSLSLOQAFSELRAQMTGENT 1068
 QY 1077 PPNSSHGTPEFVVPPLSLSLAGVPTTAAT--APVATSSPPNDISTSYQSVTYPT 1134
 DB 1069 PPNSSHGTPEFVVPPLSLSLAGVPTTAAT--APVATSSPPNDISTSYQSVTYPT 1125
 QY 1135 EGIAGVATSTGVVSGGLPIPVSESPVLSVSSITPIPVASISTSPSLQVPTSEI 1194
 DB 1126 KGIQGVTTSTGVVAGGLTTLVSEETLSSAVSSSTAPAVVYTTSGOVQAFSS-- 1183
 QY 1195 VVSTALYBVTYVANSASAGSSTAPGPKPRAVVSQQAAGSTTVGATLTVSTTSPS 1254
 DB 1184 IASTGSPGFTSTTGTGVSSVAVNAPKPTVLLQVAGNTAGVAVITVSTTTPFA 1243
 QY 1255 TAAQLSTQSSSTPTLATTVVVSASHLDTKSHSSSTGLAFSLASBSSSSGAGVSY 1314
 DB 1244 MASQPSLPGLSSSTAPLATTVVVSASHLDTKSHSSSTGLAFSLASBSSSSGAGVSY 1303
 QY 1315 ISQPGGLHPLVPSVIASTETILPQAAGPTSTPPLPVPSIPPLVOPANVPAVOQLTHS 1374
 DB 1304 VSGPGIYHPLVPSVIASTETILPQAAGPTSTPPLPVPSIPPLVOPANVPAVOQLTHS 1363
 QY 1375 QPOPALIPNOPTHCEPVDSTQPKAPGIDIDIKLEBKLSLSEHSSSGAQAHSVLE 1434
 DB 1364 QPOPALIPNOPTHCEPVDSTQPKAPGIDIDIKLEBKLSLSEHSSSGAQAHSVLE 1423
 QY 1435 SLVIESVTPGIPITTVAPSKLITSTSTCTLPNTLPLGIVALPVYVTPGVSTP--- 1491
 DB 1424 PLVVE-TVPBGITTVAPSKLITSTSTCTLPNTLPLGIVALPVYVTPGVSTP--- 1482
 QY 1492 ---VSTTSGVKGTAAPSKPLTKAPVLPVTELPAGTLESEQLPFPGPSLQSOQPLE 1548
 DB 1483 ASAPASTATGAKPGTTPPKRSLTKTYPPVGTSLAGTYCEBPPFPGPSLQSOQPLE 1542
 QY 1549 DLDQALRTLSPEKXITVTSVAVGVSMAAPTAITEAGTQPKGVSYKEGVLATSSGAV 1608
 DB 1543 DLDQALRTLSPEKXITVTSVAVGVSMAAPTAITEAGTQPKGVSYKEGVLATSSGAV 1599
 QY 1609 FKMGROVAVADGAKGKNSSEDAKSVHSESTSESSVLSASSSESTLVKPEPGIT 1668
 DB 1600 VKMGROVAVADGAKGKNSSEDAKSVHSESTSESSVLSASSSESTLVKPEPGIT 1659
 QY 1669 PGJSSDVPESAHKTASAKSDTGQPTKYGFQVTTANKVGFSSXTEDKTDTRKXG 1728
 DB 1660 SGJSLDVPSTHRTTPPEAKSSEGTQYGRGQVTTANKVGFSSXTEDKTDTRKXG 1719

QY 1729 PVASPPMDLEQAVLPVAPVPAKEKPELSEPSHLNGPSSDPAAFLSRVDGSGSPHSP 1788
 DB 1720 PVTS-PRRDSQVITPAIPAKKEPELSEPSHLNGPSSDPAAFLSRVDGSGSPHSP 1778
 QY 1789 QLSKSLIPSONLSQSLNSFNSSYMSDNESDIEDIEDIKELRLRDXHKEIOLQSRQ 1848
 DB 1779 HLCKSLPIQTLQSLNSFNSSYMSDNESDIEDIEDIKELRLRDXHKEIOLQSRQ 1838
 QY 1849 KHEIESLYTLGXKVPVAVIIPPAFLSGRRRRPRTKSGKSSRSSSLGKNSPQJGNTSG 1908
 DB 1839 KHEIESLYTLGXKVPVAVIIPPAFLSGRRRRPRTKSGKSSRSSSLGKNSPQJGNTSG 1898
 QY 1909 QSAASVHPQOOLHPNIPESGONOLQPLKPSDDNYSATSPGALSVPELSAPQ 1968
 DB 1899 QSGSVANTPQOOLHPNIPESGONOLQPLKPSDDNYSATSPGALSVPELSAPQ 1958
 QY 1969 GTSSTNTVGAITVNSQAAQAPPAWTSRKTFTDHLKLVDMARDAMNLSGRGSGHM 2028
 DB 1959 GTSSTNTVGAITVNSQAAQAPPAWTSRKTFTDHLKLVDMARDAMNLSGRGSGHM 2018
 QY 2029 NYEGPGMARFSAPOGLCISMTNLSGSAFISAASATSLGHTYSMCPPOQYGPATPFG 2088
 DB 2019 NYEGPGMARFSAPOGLCISMTNLSGSAFISAASATSLGHTYSMCPPOQYGPATPFG 2078
 QY 2089 AOWSGTGPAPOPLQGFQVGTASLQNFNISNLOKSISNPPGSNLRKT 2136
 DB 2079 TOWSGTGPAPOPLQGFQVGTASLQNFNISNLOKSISNPPGSNLRKT 2126

RESULT 8
 US-10-052-648A-38
 ; Sequence 38, Application US/10052648A
 ; Publication No. US2004000558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, David
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Colman, Steven
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: Elferman, Karen
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: MacDougall, John R.
 ; APPLICANT: Mehraban, Fuad
 ; APPLICANT: Ratturajan, Meera
 ; APPLICANT: Rothenberg, Mark
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Smithsonian, Glenda
 ; APPLICANT: Spylek, Kimberly A.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Vermet, Corine A.M.
 ; APPLICANT: Zernusen, Bryan D.
 ; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
 ; FILE REFERENCE: 21402-250 (CURA-550)
 ; CURRENT APPLICATION NUMBER: US/10/052,648A
 ; PRIOR FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: 60/262,454
 ; PRIOR FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: 60/272,920
 ; PRIOR FILING DATE: 2001-03-02
 ; PRIOR APPLICATION NUMBER: 60/284,549
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/303,229
 ; PRIOR FILING DATE: 2001-07-05
 ; PRIOR APPLICATION NUMBER: 60/262,892
 ; PRIOR FILING DATE: 2001-01-19
 ; PRIOR APPLICATION NUMBER: 60/263,605
 ; PRIOR FILING DATE: 2001-01-23
 ; PRIOR APPLICATION NUMBER: 60/269,098
 ; PRIOR FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: 60/264,159

; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 60/265,517
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/271,855
 ; PRIOR FILING DATE: 2001-02-27
 ; Remaining Prior Application data removed - See file Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 97
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 38
 ; LENGTH: 670
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-052-648A-38

Query Match 31.5%; Score 3408; DB 15; Length 670;
 Best Local Similarity 100.0%; Pred. No. 1.3e-162;
 Matches 668; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGAAGKOSTGSLPLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEYRRRRHT 60
 DB 1 MSGGAAGKOSTGSLPLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTEYRRRRHT 60
 QY 61 MDKDSRGAAATTTTTHRRFRRSVTCDSNATALELPLPLPDPSPAPAPQAPPEPH 120
 DB 61 MDKDSRGAAATTTTTHRRFRRSVTCDSNATALELPLPLPDPSPAPAPQAPPEPH 120
 QY 121 REETVNTATNSQVAAQPPAAAPGEQAVAPAPSTVSSSKDRPVSGPSLVGSKKEPP 180
 DB 121 REETVNTATNSQVAAQPPAAAPGEQAVAPAPSTVSSSKDRPVSGPSLVGSKKEPP 180
 QY 121 REETVNTATNSQVAAQPPAAAPGEQAVAPAPSTVSSSKDRPVSGPSLVGSKKEPP 180
 DB 121 REETVNTATNSQVAAQPPAAAPGEQAVAPAPSTVSSSKDRPVSGPSLVGSKKEPP 180
 QY 181 ARSSGGGSAKPEPEERSQODDIELETKAVGNSNGRFLKPDIEIGRSPFTVYGLD 240
 DB 181 ARSSGGGSAKPEPEERSQODDIELETKAVGNSNGRFLKPDIEIGRSPFTVYGLD 240
 QY 241 TETVEVAMCELODRKLTSEKRFKEAEMLKGLQHPNIVRFYDSWESTVKGKCIYLV 300
 DB 241 TETVEVAMCELODRKLTSEKRFKEAEMLKGLQHPNIVRFYDSWESTVKGKCIYLV 300
 QY 301 TELMTSGTLKTYLKRFRVMKIVLRSMCROILKGLQHLHTPTPIIHRDLKCDNIFITGP 360
 DB 301 TELMTSGTLKTYLKRFRVMKIVLRSMCROILKGLQHLHTPTPIIHRDLKCDNIFITGP 360
 QY 361 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPPEMYEKEKYDESUVVAFGCMLEMATSEY 420
 DB 361 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPPEMYEKEKYDESUVVAFGCMLEMATSEY 420
 QY 421 PYSECQNAAQIYRRVTSQVAPASPDXYAIPVEKEIIIEGCIQONKDERYSIKDLINHAFFQ 480
 DB 421 PYSECQNAAQIYRRVTSQVAPASPDXYAIPVEKEIIIEGCIQONKDERYSIKDLINHAFFQ 480
 QY 481 EETGVAVELAEEDGEEKIATLMLRIEDIKLKGKYKDNEMIEFSFLDERVPEDVAQEM 540
 DB 481 EETGVAVELAEEDGEEKIATLMLRIEDIKLKGKYKDNEMIEFSFLDERVPEDVAQEM 540
 QY 541 VESGVYECGDHKTAKAIKDEVSILKRRKROQLVREOEKKOEESLKOQVQSSASQ 600
 DB 541 VESGVYECGDHKTAKAIKDEVSILKRRKROQLVREOEKKOEESLKOQVQSSASQ 600
 QY 601 TGIKOLPSASTGIPASTTSASVSTQVPEPEADQHOOLQYQOPSIISLSDGTVDVSGG 660
 DB 601 TGIKOLPSASTGIPASTTSASVSTQVPEPEADQHOOLQYQOPSIISLSDGTVDVSGG 660
 QY 661 SSVFTESR 668
 DB 661 SSVFTESR 668

RESULT 9
 US-10-460-545-2
 ; Sequence 2, Application US/10460545
 ; Publication No. US20040005624A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Tayber, Olga
 ; TITLE OF INVENTION: 84573, A Human Protein Kinase Family
 ; FILE REFERENCE: MP102-080PLM
 ; CURRENT APPLICATION NUMBER: US/10/460,545
 ; CURRENT FILING DATE: 2003-06-12
 ; PRIOR APPLICATION NUMBER: US/60/388,031
 ; PRIOR FILING DATE: 2002-06-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1743
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-460-545-2

Query Match 22.6%; Score 2439; DB 15; Length 1743;
 Best Local Similarity 33.6%; Pred. No. 1.7e-113;
 Matches 746; Conservative 246; Mismatches 617; Indels 610; Gaps 80;

QY 27 NGSSSDSVGEKLGAAADAVTGRTEYRRRRHTMDKDSRGAAATTTTTHRRFRRSV 86
 DB 18 DGISFERNVPQ--VAATLTVEARKX-----KNSTFSASGETVERKREFRSV-- 63
 QY 87 DSNATALELPLSLPSPAPAPQAPPEPHREETVNTATNSQVAAQPPAAAPGEQ 146
 DB 64 -----EMTEDDKVAESSP-----KDERIKAAANNIPVDKLPNVLRGGQ 102
 QY 147 AVAGPADSTVPSSTSKDRPVSGPSLVGSKKEPPAPASGSGSAKPEPEERSQODDIE 206
 DB 103 EYVEEQSKSTSEISXK-----CFKENNEKEME-----EE 132
 QY 207 LETKAVGNSNGRFLKPDIEIGRSPFTVYGLDTEETVEVAMCELODRKLTSEKRFK 266
 DB 133 AEMKAVATSSGGRFLKPDIEIGRSPFTVYGLDTEETVEVAMCELODRKLTSEKRFK 192
 QY 267 EEAEMKLGLOHPNIVRFYDSWESTVKGKCIYLVTEIMTSGTLKTYLKRFRVMKIVLR 326
 DB 193 EEAEMKLGLOHPNIVRFYDSWESTVKGKCIYLVTEIMTSGTLKTYLKRFRVMKIVLR 252
 QY 327 WCROILKGLQHLHTPTPIIHRDLKCDNIFITGPVSGVKIGDLGLATLKRAFPKSVIGT 386
 DB 253 WCROILKGLQHLHTPTPIIHRDLKCDNIFITGPVSGVKIGDLGLATLKRAFPKSVIGT 312
 QY 387 PEPAPPEMYEKEKYDESUVVAFGCMLEMATSEYSEYSECQNAAQIYRRVTSQVAPASPD 446
 DB 313 PEPAPPEMYEKEKYDESUVVAFGCMLEMATSEYSEYSECQNAAQIYRRVTSQVAPASPD 372
 QY 447 VALPEVEKIIIEGCIQONKDERYSIKDLINHAFFQETGVAVELAEEDD--GEKTAIKLWL 504
 DB 373 VTDEVEVEKIIIEGCIQONKDERYSIKDLINHAFFQETGVAVELAEEDD--GEKTAIKLWL 431
 QY 505 RIBDIKLGKGYKDNEMIEFSFLDERVPEDVAQEMESGVYECGDHKTAKAIKDVVSL 564
 DB 432 -VEDPKLKGKHKMNEAIEFSFMLETTPPEVALEMKSGFPHSDSKAVAKSIRDRVTP 490
 QY 565 IKRRKROQLVREOEKKOEESLKOQVQSSASQGTIGKOLPSASTGIPASTTSASVS 624
 DB 491 IKRRKROQLVREOEKKOEESLKOQVQSSASQGTIGKOLPSASTGIPASTTSASVS 533
 QY 625 TQVPEPEADQHOOLQYQOPSIISLSDGTVDVSGGSAKPEPEERSQODDIE 679
 DB 534 TGACCEFEVDQHOVROQLQKPKQOCHSVTGDMLSEAGASVHSDPT--SSQPSVAYSS 591
 QY 680 QHEQAHSTGTPGHIIPSTV-QAQ--SQHGVYPPSSVAAQSQSQSQSSSLTGVSSQPIQ 737
 DB 592 NQ-----TMGQVAVNIPQAEVAVNPQILY-----SSQOLVIG 622
 QY 738 HPQOQGGIIOCTAPPOQVYVSLQSTSSSEATTAQVYVSLQSTSSSEATTAQVYVSLQSTSS 797
 DB 623 HYQOVSGIQRKSKLTQPIQLPLVQGO-----STVLVPIV-HVIGPVTQVQVPSPLTVQKVP 676
 QY 798 QVAP-AEPVAVAPQATQPTTLASSVDSVSHSDVAGMSDGNENV-----PSSSGRHGRT 851

Db 677 QIKPVSQVGAEOQAAALLKPLD---VRLNDQVAT---TKENSSPDNPNPGNKQORIK 729
 QY 852 TKRHYKRSVRSRSHREKTSREPKLILNNSKDRVECOLETHNRKMTFFEDLDGDNPE 911
 Db 730 QRR-----ASCPEKGTQFQVLYQVSTSGDKNVBEQQLTHNNKMTFFEDVDGDAPE 783
 QY 912 ELATIMVNDLTLAIERESFVDQVREIIEKADMLSEBVSVPEGDOGLSLOKDDYGF 971
 Db 784 DIAYWEDNDFVLESEKEKFEVEELRAIVGAQOELI--HVHATERATGVDSITVDSNSQ 841
 QY 972 SGQCKLESEFKQPIIPASSMPOQIGIPTSLSLQVHSHAKRTIVPVPESRLRESKVPSE 1031
 Db 842 TGSS-----BOYQINSTSTQTSNESAPQS---SPVGNMRFCC----- 874
 QY 1032 ITDVAASIAQSPGMNLSHSAASLSLQAFSELRAQCTEGEPNAPMFHSTGTPFPVP 1091
 Db 875 INQIRREHROS-----PFSLOH----- 892
 QY 1092 PFLSSINGVPTTAATAVPATSPPN-DISTSVIOSEVTVPTREGIAGVATSTGVVTSQ 1150
 Db 893 ---SMSAVP---GRHPLFSPKNTSNKEISRDLTLIENNPCRALFTSKSEHKDYVDG 944
 QY 1151 GLPIPVSESPVLSVSSITIPAVSISTSPSLQVPTSELY---VSTALYPSVTY 1207
 Db 945 ---KISECASVETKQPIALYQVEDNRQIMAPVYNSSYSTSV 984
 QY 1208 SATSASAGSTATPPKPPAVNSQQAAGSTVGATLTSTSTTSPSTASOLSTQLSST 1267
 Db 985 RAYVAEBCGLTKQASIFIPVPCHQ-----TASQADALMHPESTQ- 1026
 QY 1268 STPTLATVAVSAHSLDKTSHSSTTGLAFSLAPSSSSSPAGVSVYISQPGGLHPLVP 1327
 Db 1027 ---TSGNSLTLTAFQCKPQTLISVQCPAMDAERISQEG-----E 1061
 QY 1328 SVLASTPILPOAGPTSTPLLPQVSPPLVQVPAVNPVQOTLIHQPOBALHPQPT 1387
 Db 1062 TYVNTASSPKTYVITPTGLE-----PTTLQ- 1090
 QY 1388 HCEVDSDTOPKARGIDD-IKTLLEKLS--LFSHSSSGAQHASVLSLSTSLVSTVTPG 1445
 Db 1091 TVLESDBERPKLEFANRIKITDEKRLNLYQHSIS-----SIYESQ- 1135
 QY 1446 IPTTAAPSKLSTSTSTCLPPTNLPGLVALPVTPVT--PQVSTPVST---TTS 1497
 Db 1136 ---KDTOSIDSPFSSABDTLSCPTEVIALSHGIGIKSPQSPNFQGTGS 1183
 QY 1498 GVKPGTAPSKPP-----LTKAPVLPVGTLPBAGLPSB-QLP--PFGPSLTQS 1543
 Db 1184 KILSNVAASOPANISVFKRDLNVITVP-----SELCLHMSSDASLPDPEAYPAVVS 1238
 QY 1544 QOPLDLDLQRLRTLSPEXITVTSVAVPVSMAPATLTAAGTQPOKQVSVKGPVLATS 1603
 Db 1239 ---GAIHQVTEEMSAIAP---DPILTRSTADTRALNCK----- 1279
 QY 1604 SGAGVPMKGFQVVAADGAQKKGKSEDAKSVHESSTSESVLS-SSSPSTLVKPE 1662
 Db 1280 AMGSGFQGRFQVITTPQ--QOSAKMTSFGIEHISVSESTNHSSEAFITAKSQVLE 1337
 QY 1663 PNGITTPGISDVPEASHKTTASEAKSDTQPTKVGFRQVTTTANKVGRSVKTEDKIT 1722
 Db 1338 P-----ATQNPKTSPSYEKLQ-ALQ 1356
 QY 1723 DTKKEG---PVASPPMDLEQAV---LPVAVPKKPE-----LSEPSH 1761
 Db 1357 ETCKEKGVKQGDNLFSFACETDVSVTPKEKEFHISATSSMOGSELKREIL 1416
 QY 1762 NG---FSSDPE--AATLSRDVDDSG-----SPHS---PHQJSSKSLPQNLISQSLNS 1807
 Db 1417 TACKQSSSSEFASIA-----GSGKSVAKTGPESNOCLPH-----EBOVAAQTOSS 1465
 QY 1808 F--NSSVMSDNDSDIEDEDLKLELRIRDKHKEIODOSRQKHEIESIYTLQGVAPPA 1865

Db 1466 FVSPSSPMSSDSEIEDEDLKVELQRLREKHIOEVNVLQTOCKNEIQLERYLRISIKDS 1525
 QY 1466 -----VILPPAATLSRRRRPTSKSGSKSSRSSLSQKSPQSLNSGSAABVTHPQ 1919
 Db 1526 KTQSTELPLPABP-----RRPNSFKSKLSR-----FQS 1555
 QY 1920 TLHPNGNIPESGQQLLQPLKPSFSSDNLXSAFTSDGALSVPSLAPQGTSTNTVAT 1979
 Db 1556 LTH-----VDN-----GLVADPLCIVE 1572
 QY 1980 VNSQAQAQAPAMTSSKGTFTDHLKLVNWARDAM-----NLGRRGSKGHMNYEG 2032
 Db 1573 SNAASCOQSP-----ASKKGMFTDHLKLVDMTKEAVGNLSIKPSLNLQKSOHKLFTEN 1628
 QY 2033 PGMAKFSAPQOL-----CISMNSNLGSAPIASAATSLGHFTKSMCPPOQYGP 2083
 Db 1629 MNKYSB--NTPSTNGYSTWISSLSQIRGAVPTSLPQSLSPSFGP---PLSSYMPHYCQ 1684
 QY 2084 ---ATPFGAOWSGTGPAPQPL-----GQFQPVGTASLQVFNISNLQKSTSNPPG 2130
 Db 1685 YNAVAGGYVQVWVGISGTTQGSVVIPOSGGCFQP--GNMMAFTSSVQNPATIFPG 1741
 RESULT 10
 US-10-433-794-17
 ; Sequence 17, Application US/10433794
 ; Publication No. US20040077044A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YUE, Henry; DING, Li;
 ; APPLICANT: LAB, Ereeti G.; GRIFFIN, Jennifer A.;
 ; APPLICANT: GURURAJAN, Rajagopal; BAUGHN, Mariah R.;
 ; APPLICANT: ISON, Craig H.; RAMKUMAR, Jayalaksmi;
 ; APPLICANT: TRIBOULEY, Catherine M.; SWARNAKAR, Anita;
 ; APPLICANT: BURFORD, Neil; BANDMAN, Olga;
 ; APPLICANT: THORNTON, Michael; KHAN, Farrah A.;
 ; APPLICANT: WALIA, Narinder K.; NGUYEN, Daniel B.;
 ; APPLICANT: ELIOTT, Vicki S.; XU, Yuming;
 ; APPLICANT: LU, Yan; HAPALIA, April J.A.;
 ; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
 ; APPLICANT: ARVIZU, Chandra S.; FORSTHE, Ian J.
 ; TITLE OF INVENTION: KINASES AND PHOSPHATASES
 ; FILE REFERENCE: PI-0311 USN
 ; CURRENT APPLICATION NUMBER: US/10/433,794
 ; PRIOR FILING DATE: 2003-06-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/47431
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: US 60/254,034
 ; PRIOR FILING DATE: 2000-12-06
 ; PRIOR APPLICATION NUMBER: US 60/251,814
 ; PRIOR FILING DATE: 2000-12-07
 ; PRIOR APPLICATION NUMBER: US 60/255,756
 ; PRIOR FILING DATE: 2000-12-14
 ; PRIOR APPLICATION NUMBER: US 60/256,172
 ; PRIOR FILING DATE: 2000-12-15
 ; PRIOR APPLICATION NUMBER: US 60/257,416
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 60/260,912
 ; PRIOR FILING DATE: 2001-01-10
 ; PRIOR APPLICATION NUMBER: US 60/264,644
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: US 60/266,017
 ; PRIOR FILING DATE: 2001-02-02
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 17
 ; LENGTH: 1345
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 7482044CD1
 US-10-433-794-17

Query Match 19.7%; Score 2130.5; DB 16; Length 1345;
 Best Local Similarity 38.8%; Pred. No. 3.6e-98; Indels 475; Gaps 36;
 Matches 537; Conservative 114; Mismatches 258; Indels 475; Gaps 36;
 Db RFRFRSVICDSNATALELPGILPLSLPQPSIPAAVPGAPPEPHREFT-----VTAT 128
 78 RFRFRSVICDSNATALELPGILPLSLPQPSIPAAVPGAPPEPHREFT-----VTAT 128
 26 RFRFRSVICDSNATALELPGILPLSLPQPSIPAAVPGAPPEPHREFT-----VTAT 128
 Db RFRFRSVICDSNATALELPGILPLSLPQPSIPAAVPGAPPEPHREFT-----VTAT 128
 129 ATGCAVAAQPPAA-----AAGEQAVAGPAPSTVSSSKRPVSQPSLVSKKEPP-- 180
 75 ARGRPAAPAPAAVAGPAGAPADAGPEP-----VGIOEPGDPPTAAAVETAPADGG 128
 181 -----ARSGGGGSAAK-EP-----QEESSOQODDIEELETAKVMSNDGRFLKDE 226
 129 PREBAATVRKEDGAAEAAPPEPRTTRDEPEDEEDDL--KAVATSDGRFLKDE 186
 227 IGRSFTVYKGLDTEETTVAVMCELODRKLTKSERQFKEEAMLKGLOPHNIVREYDS 286
 187 IGRSFTVYKGLDTEETTVAVMCELODRKLTKSERQFKEEAMLKGLOPHNIVREYDS 246
 287 WESTVKKKCIIVVTEMTSGTLKTYLKRKWKIKYLRBWCROILKGLQPHNTRPPI 346
 247 WESSAKKRCIIVVTEMTSGTLKTYLKRKWKIKYLRBWCROILKGLQPHNTRPPI 306
 347 HRDLKCNIFITGPTGSVKIGDLGLATLTKRASFASVIGTPEFAPAMEEKKYDESVDVY 406
 307 HRDLKCNIFITGPTGSVKIGDLGLATLTKRASFASVIGTPEFAPAMEEKKYDESVDVY 366
 407 AFQMCMLMATSEYYSSECONAAQIYRVTSGVKPSFPAKVAIPVKEIIEGCIQNKDE 466
 367 AFQMCMLMATSEYYSSECONAAQIYRVTSGVKPSFPAKVAIPVKEIIEGCIQNKDE 426
 467 RYSLKDLINAFPEELGVAVELAEEDGER--IAITLMRIEIKLKVKYKQNEAIEF 524
 427 RYSLKDLINAFPEELGVAVELAEEDGER--IAITLMRIEIKLKVKYKQNEAIEF 484
 525 SFDLERDPEVAOEMESGVVCEGDKHTAKAIKDVSLIKRREOROLVREOEKKQ 584
 485 TFDLEKTPPEVAOEMESGVVCEGDKHTAKAIKDVSLIKRREOROLVREOEKKQ 542
 585 EESSLKQOEVSSASQGTGINKQPSASTGIPASTTSASVSTOV-----EPPEADQHQ 638
 543 DVGS-----PKARGPPVPLQVTVYHAQAGQPPPEPEPEADQH- 583
 639 QLOVQOQSI-----SVLSDGVDSGQSSSVTRESVSSQOQVSGSQHEQAHSTG- 688
 584 ---LLPPLTPSATSLASDSTFDGQGSQVYSDQ--SSQGSVMLGSLADAPSPAQCVCS 639
 689 ---TVPGHISTVQAQSQPHG-----VYPP 710
 640 PVSEGEVLPQSLPS-LGAYQPPTRAAPGLPVGSVPAPACPPSLQOHPPDPAASFAPVLP 698
 711 SSVAQGSQGSQPS-----SSSLTGV 730
 699 PSTMPPTGPGQAPAPGQPPPLAQPTPLPVLAQPVPLQPVPLPPLPPLAPASQVAP 758
 731 SSSEPIQHPOQO-QGIIQO-----TAPQOQTVQ 756
 759 AQLKPLQMPQAPLQPLAQVPPQMPPIVPPITPLAGIDGLPALPDLPLTATVPPMPPP 818
 757 Y----- 757
 819 YFSRAVILPSLAALPPLASAPLPLQAVLPHRPGAPLAMPCKRTIVNAPATIPLLAVAP 878
 758 ---SLQSTSSSATTAAQVVSQOQ----- 778
 879 GVALSIHSAVAQUPGQPVVPAFAFPQMAPTDVPPSPHTVQNMBAITPQPALPPOPTLPP 938
 779 --- 778
 939 QPVLPPQPTLPPQPVLPQPTLPPQPVLPQPMPLPQCVLPQPALVPPPLQPHLEQ 998
 779 ---APQVLPQ----- 785

Db AAPATPSSQIILGHAPYAVDAQVPTVPPPAVAPLPPLEVLLPAPAPLPPQFPSS 1058
 786 ---VSAGKOS--TQVSVQVAP-----EPVAVQAPQATQPTTLA 819
 1059 LATVASASQSVPTQPTLPPPPANPLPGPGIAPSPPTVQLTVEPVOEBSAQDKPPGPP 1118
 820 SSVDS-AHSDVASG--MSDGENVPPSSGHRGRTTKRHRKSVRSRREKTSRPLRT 876
 1119 QSCSTGSGSDVTSKGLSCEGA-FGGKLBGRARAKHRSTRARSROEARSPLRT 1177
 877 LANSNKGDRVVECOLETHNRKMTFKFDLDGNDPBEIATIMVANNFIALERESFVDQVR 936
 1178 LANCNGDKMVEQLETHNRKMTFKFDLDGNDPBEIATIMVANNFIALERESFVDQVR 1237
 937 EITKADMEISEVSVPEBDOGLE---SLQKDDYGFSGSKLBEFQOPTPASSMPQ 992
 1238 DVMDKXEDMLSEPTDDRGSDPSTSPHLSCTGIG---TGESSROSQANAPV---YQ 1288
 993 QIGIPSSILQVYHSAGRFIVSPVPSRLRESKVPFSEITDVAASTAOSPGKMLSHSA 1052
 1289 Q-----NVHTGKRWITICPVAEHAPBA---PSSPPLPLSSLPCLPFLRMSCA 1335
 1053 SLSL 1056
 1336 SVLA 1339

RESULT 11
 US-10-052-648A-37
 ; Sequence 37, Application US/10052648A
 ; Publication No. US20040005558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, David
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Coleman, Steven
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Macdougall, John R.
 ; APPLICANT: Mehraban, Fuad
 ; APPLICANT: Pattnajain, Meera
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Smltheon, Glenda
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Vernet, Corine A.M.
 ; APPLICANT: Zephusen, Bryan D.
 ; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
 ; FILE REFERENCE: 21402-250 (CURA-550)
 ; CURRENT APPLICATION NUMBER: US/10/052,648A
 ; CURRENT FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: 60/262,454
 ; PRIOR FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: 60/272,920
 ; PRIOR FILING DATE: 2001-03-02
 ; PRIOR APPLICATION NUMBER: 60/284,549
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/303,229
 ; PRIOR FILING DATE: 2001-07-05
 ; PRIOR APPLICATION NUMBER: 60/262,892
 ; PRIOR FILING DATE: 2001-01-19
 ; PRIOR APPLICATION NUMBER: 60/263,605
 ; PRIOR FILING DATE: 2001-01-23
 ; PRIOR APPLICATION NUMBER: 60/269,098
 ; PRIOR FILING DATE: 2001-02-15
 ; PRIOR APPLICATION NUMBER: 60/264,159
 ; PRIOR FILING DATE: 2001-01-25

```

; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See file wrapper or PAM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 37
; LENGTH: 1231
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-052-648A-37

Query Match      19.7%; Score 2126; DB 15; Length 1231;
Best Local Similarity 31.9%; Pred. No. 5,4e-98;
Matches 612; Conservative 191; Mismatches 393; Indels 724; Gaps 58;

QY 10 GSTPGSLFLSPAPAPKXSSSSSVGEKLGAAADAVTGTEEYRRRRHTMDKSRGAA 69
DB 2 SQTADLALRPPLPLGTAGQ-----PRLGPP-----RRARFSGK----- 37
QY 70 ATTTTTHRRPFRSVIDCSNATALELPLSLPQPSIPAPVPOSAPPEPHEETVATA 129
DB 38 AEPFRSSRLSRKSSVLDGLSSWSLPASP--APDPDP--PDSAGGPAR----- 84
QY 130 TSQVAAQPPAABAGEAVAGPAPSTVPSSTKDRPVASQPSLVGSKKEPPPARSGSGGS 189
DB 85 -----SPPPSKEPEGTWTEGAP--VXAADSAKPELPDSAVPGSRP----- 127
QY 190 AKEPQOE--ERSQODDIEELIETAVGMSNDGRFLKPDIEIGRSFRTVYGLDTETTE 246
DB 128 LRVEAVALERRRRQEEKEDMETQAVATSPGRYKLFEDIEIGRSFRTVYGLDTETTE 187
QY 247 VAMELODRKLTKEBROFKEAEMLKGLCHPNIVRFDSMESVYKGCIVLTETMTS 306
DB 188 VAMELODRKLTKEBROFKEAEMLKGLCHPNIVRFDSMESVYKGCIVLTETMTS 247
QY 307 GTLTKYLKRFKVMKIKVLRSCWQIILKGLQFLATRTPEIIRHDKCDNIFITPGTSYKI 366
DB 248 GTLTKYLKRFKVMKIKVLRSCWQIILKGLQFLATRTPEIIRHDKCDNIFITPGTSYKI 307
QY 367 GDLGLATLKASPAKSVIGTEPFMAPEMYEKEKYDESVDVYAFGCMLEMATSEYPSYEQ 426
DB 308 GDLGLATLKASPAKSVIGTEPFMAPEMYEKEKYDESVDVYAFGCMLEMATSEYPSYEQ 367
QY 427 NAAQIVRVNVSQVAPSPDKVAIPEVKEIEGCIKONDERYSIKDLNHFPEDEGVR 486
DB 368 NAAQIVRVNVSQVAPSPDKVAIPEVKEIEGCIKONDERYSIKDLNHFPEDEGVR 427
QY 487 VELAEEDGEEKIALKWLRIEDIKLKGYKUNBAIESFPLERDVEDVYAEWVESGYV 546
DB 428 VELAEEDGEEKIALKWLRIEDIKLKGYKUNBAIESFPLERDVEDVYAEWVESGYV 546
QY 547 CEGHKTAKAIKDRVSLIKRKEQKQVREEOEKKKQEESSLKQVEQSSASQTKQL 606
DB 487 CEGHKTAKAIKDRVSLIKRKEQKQVREEOEKKKQEESSLKQVEQSSASQTKQL 606
QY 607 PSATGIPPTASTASVSTQY--EPREPADQHQLQYQPSI--SVLSDGTVSGGSS 662
DB 522 P-PPGPAPATVMAFGPPSVFPPEPEPADQHQLFLFRASSYSTTSDCTGYLSSS 580
QY 663 VFTSRVSSQQTAVGYSQHEQAHSTGTVPCHI PSTVQAQSQPHGYVPPSSVAQSQSQGP 722
DB 581 GFLDASDPAALQ-----PGGVPSL--ASHCLPESA----- 610
QY 723 SSSSLTVGVSSSQPIQHPOQGGIQTAPDQQTVOYISLSTSTSEATTAGPVSOPOAV 782
DB 611 -----FALS----- 614
QY 783 LPQVSAGKSTQGVQVAPAPVAVAPQATQPTTLASVDSASHDVSGMDGNENVS 842
DB 615 IPRGPGSDFSPG-----DSVYSDASAGSLDVGEGM-G 646

```

```

QY 843 SSGRHEGRTTKRYKRSVRSRSHKTSRPKILNLVSNKGRVVECOLETHNRKMYTEK 902
DB 647 QMRPPGRNLR-----RPRSR-----LWTVSDQNDVRVECOLQTHNRYMYTFR 692
QY 903 FDLDDGNDEEIALTMVNDLILAIERESFVDQREILKXKADMEISEDVVPEBDQGLS 962
DB 693 FDLDDGNDEEIALTMVNDLILAIERESFVDQREILKXKADMEISEDVVPEBDQGLS 962
QY 963 LOGKDDVGFSGQKLEGEFKQPIPASMPQOIGPISLTQVYHSAGRRTVSPESRL 1022
DB 753 QF-----EPALPLPPLPDPSPBEQ-----SSTS 779
QY 1023 RESKVPSEITDVAASTAOSPGCNLHSHASISLQAFSELRRAQMTGENTAPPNFSH 1082
DB 780 LEHRMSTAFSTSSSPCTPLSPG--NPPSPGPIPSPCIF-----PITSPP--CH 825
QY 1083 TGPTFPVPPPLSLIACVPTTAATAVAPATSSPNDISTSVIOSEVTVPEESGIAGVAT 1142
DB 826 PSPS--PSPSISQVSSNPSHPPTSPSPSSSTPE----- 859
QY 1143 STGVVTSGLPPIPVSESPVSSVSSITIPAVVISITSPSLQVPTSEIVVSTALY 1202
DB 860 -----FPV-PLSQCP-----WSLPTTSPPTSP--CQVTLSSPFF- 894
QY 1203 PSVTSATASAGSGTATPGPKPPAVVSGQAAGSTVGATLTVSTTSPSTASQLSIQ 1262
DB 895 -----PFC-----PSTSPSTTAADLS 913
QY 1263 LSSSTPTPLAETVYVSAHSLDKTSHSTTGIAFLSLAPSSSSPAGAVSSYSIQPGLH 1322
DB 914 LAS-----AFSLA----- 921
QY 1323 PLVPSVIASTPILPQAAGFTSPILPQVSPILPVQVAVNPAVQOTLIHSCQPALP 1382
DB 922 -----VMTVAAGSL--SPSPGLS 938
QY 1383 NQPHTHCEVSDSTQPPAPGIDIKTLEKRLSLFSEHSSGAGHASVLETSVISTV 1442
DB 939 QSP----- 941
QY 1443 TPGLPTTAAVAPSKILTSTTCLPPTNLPLGTVALPVTVPVTPQGVSTPVTTSYKXPG 1502
DB 942 -----PA-----PSPPLP--SLPPLPVPAPGQSP--SPHTAEV-- 973
QY 1503 TAPSKPLTGAAPVLPVGTLPAGTLFSEQLPPFPSPSLTQSQPLEDLDLAKRTLSPEX 1562
DB 974 -SEASPP-----PARLPGE----- 987
QY 1563 ITVTSAGVPVSMAPTAITAGTQPOKGVQVKEGVLATSSGAVFERKGRFOYVVAADG 1622
DB 988 -----ARLAPIS-----ERKQPL-----VGRFQVT----- 1008
QY 1623 AQKEGKXKSEDAKSVHEESSTSESVSSSPESTLVKPPBNGLTIGJSSDVESAHKT 1682
DB 1009 -----SSKPAEPLPQPT--SPTLSSG--PKRSTPLTSESDTDESA-- 1048
QY 1683 TASEAKSDTQOPTVGRFOYTTANKVGRSVSKTEKIDTKKEGVPVAPPFMDLQAV 1742
DB 1049 -----GGGP-----ETRALAESDR----- 1063
QY 1743 LPAPVTPKKEKREISEPSHLNGSSDPEAFLSRDVDGSGSPH--SPHOLSSKSLPSQN 1799
DB 1064 -----AAEGLGAGVEEGDDGK--EPYVGSGSPQPLSHSPVWVN 1100
QY 1800 LSQSLNSPNSSTVSSDNESIDIBEDKLEHRLRDKHLEIDLOQRKHEISLTKL 1859
DB 1101 YSYS-----SLCJSSESESGEDEFWAELOSLRQHLSEVETLQTLQKKEIDLSRL 1155
QY 1860 GKVPAPVLIIPAAPLSGRRRRPTKSKGS--KSSRSSSLGNKSPQLSG-----NLGQSAAS 1913
DB 1156 GKQPPPGIVAPAAALSLGRQR--LSKSGPITSRRNLSIQRSPPPGIMRRNRSLSSGSG 1213

```



```

RESULT 12
US-10-196-935A-4
: Sequence 4, Application US/10196935A
: Publication No. US20030082720A1
: GENERAL INFORMATION:
: APPLICANT: Lifton, Richard P
: APPLICANT: Wilson, Frederick H
: APPLICANT: Choate, Keith
: APPLICANT: Ishikawa, Kazuhiko
: APPLICANT: Nelson-Williams, Carole
: TITLE OF INVENTION: COMPOSITIONS METHODS AND KITS RELATING TO TREATING AND DIAGNOSING
: FILE OF INVENTION: HYPERTEXTENSION
: FILE REFERENCE: 044574-5113
: CURRENT APPLICATION NUMBER: US/10/196,935A
: PRIOR FILING DATE: 2002-10-25
: PRIOR APPLICATION NUMBER: US 60/306,084
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 1243
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-196-935A-4

Query Match      19.7%; Score 2126; DB 14; Length 1243;
Best Local Similarity 31.9%; Pred. No. 5.5e-98;
Matches 612; Conservative 191; Mismatches 393; Indels 724; Gaps 58;

QY 10 SSTFGLFLSPAPAPAKNGSSDSSVGKLGAAAADAVTGRTEBYRRRRHTMDXSGAA 69
DB 14 SQTADIALRPPPLGTAGQ-----PRLGPP-----RRRRRSGK----- 49
QY 70 ATTTTTHRRFRSRVICDSNATLELRLPLSLPQSPAPVPOSAPPEPRRETWTATA 129
DB 50 AEPFRSSRLSRSSVGLGLSSWSLPASP--ADPDPP--PDSAGGPAP----- 96
QY 130 TSQVAAQPPAAAAGEQAVAGPASTVPSSTKORPVASQPSLVGSKPEPPPARSGSGGS 189
DB 97 -----SPPSSSKPEPEGTWTEGAP--VKAADSAARPELPDSAVGPGSEPP----- 139
QY 190 AKERQEE--ERSQOQDDIELEETKAVGNSDREFLKDIEIGRSFKTVYKGLDTEETVE 246
DB 140 LRVEAVALERRRRRQEEKEDMETQAVATSPDGRYLKFDIEIGRSFKTVYKGLDTEETVE 199
QY 247 VAWHELODRKLTKEERQRFKEBAEMKGLQHPNIVRFRDWSSTVYKGCIVLTELMTS 306
DB 200 VAWCELOTRKLSRAERQRFSEVEMLKGLQHPNIVRFRDWSKSVLRQVCIVLTELMTS 259
QY 307 GTLKTYLKRFKVMKIKYLRSCWQIILKQLOPLHTPTPIIHRDLKCNIFITPTGSVKI 366
DB 260 GTLKTYLKRFKVMKIKYLRSCWQIILKQLOPLHTPTPIIHRDLKCNIFITPTGSVKI 319
QY 367 GDLGLATLKRAKSPKSVIIGTEPFMAPMEYBEKYDESVDVAFGCMLEMATSEYSECQ 426
DB 320 GDLGLATLKRAKSPKSVIIGTEPFMAPMEYBEKYDESVDVAFGCMLEMATSEYSECQ 379
QY 427 NAAQIVRRVNSGVPAFPAFKAIPVEVEIIEGCIROKDEKYSIKDLINAFPEBEYGR 486
DB 380 NAAQIVRRVNSGVPAFPAFKAIPVEVEIIEGCIROKDEKYSIKDLINAFPEBEYGR 439
QY 487 VELAEDDGEKIAIKLWIRIDIKKLGKYKONENAIFFSPDLERDVEDVAVQEVESGYV 546
DB 440 VELAEDDGEKIAIKLWIRIDIKKLGKYKONENAIFFSPDLERDVEDVAVQEVESGYV 498
QY 547 CEGDHKTMATAIKORVSLIKRREORQLVREDEQKKQOESLSKQVQESASQGTIKOL 606
DB 499 CEGDHKTMATAIKORVSLIKRREORQLVREDEQKKQOESLSKQVQESASQGTIKOL 533
QY 607 PSASTGPIFSTISASSTQV--EPREPADQHQOLQYQOPST--SVLSGTVDSGGSS 662
DB 534 P-DEPGPPATVVPAPGPPSVFPPEBEPBEADQHQPFLFRHASYSTTSOCEOTDGYLSSS 592

```

```

QY 663 VETESRVSSQOQTVSYSGHQHQAHTGTVPGRHIBSTVQAQSPHGVPSSVAQGSQSGQ 722
DB 593 GFLDASDPALQP-----PGGVPSL--ASHCLPESA----- 622
QY 723 SSSSLTVGSSSQPIQHPOQOQGIQOTAPRQQTQVSYLSQTSSTSEATTAQVPSQAPQV 782
DB 623 -----FALS----- 626
QY 783 LPQVSAGKOSTQGVQVAPAPVAVAPQATQPTTLASSVDNSAHSDVAGSGDGENVPS 842
DB 627 IPRGSGSDFSFG-----DSYASDAAGSLDVEGEM--G 658
QY 843 SSGRHEGRTTKRHRKRSVSRREKTSRPKLRILNYSNKNDRVVECOLETHNRMVTFK 902
DB 659 QMRPPGGRNLKR-----RPRSR-----LRVTSVDQNDRVVECOLQTHNSKMTYFR 704
QY 903 FDLDDGNDPEELATIMVNNDFILATERESFVDQVEIIEKADENLSEDSVPEPGQGLS 962
DB 705 FDLDDGNDPEELATIMVNNDFILATERESFVDQVEIIEKADENLSEDSVPEPGQGLS 764
QY 963 LQCKDYGFGSSQKLBEFFKQPIPASSMPQOIGIPTSSLTQVHASGRRTVSPVESRL 1022
DB 765 QE-----EPADLPALPVPLPDPNSNELO-----SSTS 791
QY 1023 RESKVPSEITDTVAASTAOSPQGNLHSHASLSLQAFSLRRAQMTGPNTPAPNPSH 1082
DB 792 LEHRSWTAFTSSSPGTPPLSPG--NPPSPGTPISPGPIF-----DITSPP--CH 837
QY 1083 TGPTFPVPPFLSSIAQVPTTAATAAPVATSPSPNDISTVIGSEVTVPTREGIAGVAT 1142
DB 838 PSPS--PSPSISQVSSNPSHPPTSLPSSSTPE----- 871
QY 1143 STGVTSGLPIPIPVSSPVLSSVSSITIPAVVISITTSLOVPTSTETIVVSTALY 1202
DB 872 -----FPV--PLSOP-----WSSLPTSPPTSPPT--CSQVTLSSSEFF-- 906
QY 1203 PSVTASATASAGSSTTPGPKPPAVVSSQAAGSTTGATLTSVTTSPSTASQSLQ 1262
DB 907 -----PDC-----PSTSPSTTAAPLSS 925
QY 1263 LSSSTPTPLAETVVSASHLIDKTSSTTGALFSLAPSSSSPGAGVSSYISQPGALH 1322
DB 926 LAS-----AFSLA----- 933
QY 1323 PLVPSVIASTPLIPQAGPTSTPLPLQVPSIPPLVQVAVNAVAVOQTLHSQOPALLP 1382
DB 934 -----WMTVAQSLT--SPSPGLS 950
QY 1383 NQPHTHGPEVSDPTQKAPGIDIKLEKLRSLFSEHSSSGAQAASVLSLTSVISTV 1442
DB 951 QSP----- 953
QY 1443 TPGIPTTAAPSKLSTSTSTCLPNTLPLGTVALPVPVTPGQVSTPVSTTSGVKG 1502
DB 954 -----PA-----SLPFPVAPGQSGSP--SPHTAIVE-- 985
QY 1503 TABSKPPLTAPVLPVSTELPAGTLPEBQPLPPPGBSLTQSQPLDLIDQRLRTLSPEX 1562
DB 986 -SEASPP-----PARPLPGE----- 999
QY 1563 ITVTSAGVPVSMAPTAITBAGTOPQKGVQVKEGVLATSSGAGVFKMGRFOVSAADG 1622
DB 1000 -ARLAPIS-----BEGKQUL-----VGRFOYT----- 1020
QY 1623 AOKRGNKSEDAKSVHPESSSTSSSVLSSSPSTLVKPPPNGITTPGISDVPEASHT 1682
DB 1021 -----SKKEPAPLPLQPT--SPTLSSG-----PKPSTPQLTSSSDTDESAA-- 1060
QY 1683 TASEAKSDTQOPTKVGAFQVTTANKVGRSVSTEDKIKITDTKKEGVYASPPMDLQAV 1742
DB 1061 -----GGGP-----ETREALAESDR----- 1075
QY 1743 LPAVTPKKEPELSEPSHLNGPSSDPEAFLSRDVDGSGSPH--SPHQLSKSLPSQN 1799

```


LENGTH: 1251
 TYPE: PRM
 ORGANISM: Homo sapiens
 US-10-114-270-80

Query Match 18.9%; Score 2042.5; DB 12; Length 1251;
 Best Local Similarity 30.0%; Pred. No. 8.6e-94;
 Matches 605; Conservative 180; Mismatches 430; Indels 803; Gaps 56;

10 SSTPSLSPAPAPKXGSSSDSVGEKLGAAADANTGRTTEVRRRHMDKSRRA 69
 2 SQTEDLRLRPPLPLGTAGQ-----PRLGPP-----RRARRSGK----- 37
 70 ATTTTHERFRFRSVICDSNATALEPLPLSLPPOSIPAAVPOSAPREPREETVATA 129
 38 AEPFRSRLSRSSVDGLISWMLPASP--APDPPP--PDSAGCPAR----- 84
 130 TSQVAAQPPAAAPCEQAVAGAPSTVPSSTSKDPVPSQPSLVGSKKEPPPARSGSGGS 189
 85 -----SPPSSKEPPEGTWTEGAP--VKAADSARPELPDASVCGSREP----- 127
 190 AKERPE--ESSOODDIELEETKAVGMSNDGRFLKPDIEIGSGSKTVYKGLDTEETVE 246
 128 LRVEAVALERREKEDEMETQAVATSPGRYUKFEDIEIGSGSKTVYKGLDTEETVE 187
 247 VAWCELO--DEKLTSEKREKFEAEMLKGLQHPIVAFYDSESTVKGKCIYLVEIM 304
 188 VAWCELOVTRKLSRAERQRFSEVEMLKGLQHPIVAFYDSESTVKGKCIYLVEIM 247
 305 TSGTLTKYTKFKYKWKIKVLSWCRQIIKGLQFLHTPTPIIHRDKCDNIFFTGPGSV 364
 248 TSGTLTKYTKYRFRMKRPVLOKWSRQILKGLHPLHSRVPPIIHRDKCDNIFFTGPGSV 307
 365 KIGLGLATLKRASFAXSVIGTPEFMAPEWEEKYDSDVYARGMCLMWTSEYPSSE 424
 308 KIGLGLATLKRASFAXSVIGTPEFMAPEWEEKYDSDVYARGMCLMWTSEYPSSE 367
 425 CONAAQIYRRTSVGVKPAFDKVAIPEVKEIIIEGICRQNKDERYSIKOLLNHAFOEETG 484
 368 CONAAQIYRRTSVGVKPAFDKVAIPEVKEIIIEGICRQNKDERYSIKOLLNHAFOEETG 426
 485 VRVLAIEDDEDEKAIKIMLIEDIKKUKKYYKNEALIESFDLERDVPEDVAOEYVSG 544
 427 VHVLAIEDDEDEKAIKIMLIEDIKKUKKYYKNEALIESFDLERDVPEDVAOEYVSG 485
 545 YVCEGDHKTMAKATKDRVSLTKRREKQOLVREGEKKOESESLKQOVRESSASQGTGK 604
 486 LVCEADYQPVARAVRERVAALQKREKLRKARVAAIQKREKLRKARVAAIQKREKLRKAR 537
 605 QLPASASTGIPASTTASVSTOV--EPREPADQHOOLQYQPSISVLSDGTVDGSGGS 661
 538 ALP-PEPGPPATVPMABGPPSPVPEPEBEADQHPFLFRHASY----- 583
 662 SVTFRESVSSQIVSYSGSHQMASTGTVPCHISTVQAQSQPRGVTPSSVAGQSGQGG 721
 584 -----SS 585
 722 PSSSLTGVSSSQPIQHPQOQGIQOTAPQQTQVQYSLSTSTSSSEATTQAQVSPQAPQ 781
 586 PSNS-----YA----- 591
 782 VLPQVSAKOSTQGVSAVAPAEVAVAPQATOPTTLTASVDSAHSDVASGMSDGNENVP 841
 592 -----SDAAGLSIVGEGM- 605
 842 SSSGRHEGRTTKHRYKSVSRNHEK---TSRPKLRI-LVNSKNGRVVBCOLETNNRK 897
 606 GQMRPRGRMLRRFRRLRVTLSQHDEGLTFSPLTLNLTQVSDQNDRVVECOLQTHNSK 665
 898 MTFKFLDQDNPEELITIMVNDPILAIRESVVDQVRELIIEKADMLSEDSVVEEGD 957
 666 MTFRFDLDQSDPEELIAAVVINEFILPSEBDGFLIRIRIQLQVETTLKRD----- 717

958 QGLESLQKDDYGFSGSQKLEGEFKQPIRASSMPOQIGIPRSSLTQVYHSAGRRTVSPV 1017
 718 ----- 717
 1018 PESHRESKVPSEITDVAASTAOSPGMNLSHSASLSLQOAFSELRAQMTGPTAP 1077
 718 -----PLTSP 722
 1078 PNFHTGTFPVVPPFLSSINGVTTAAATAPVATSSPPNDISTSVIQSEVYVTEBGI 1137
 723 P-CHPSS--PFSPLSSQVSNPSPHPTSSPLPSSSTPE----- 759
 1138 AGVATSGVWTSGLPIPVESPVLSVSSITIPAVVSISTSPSLQVPTSEIYVS 1197
 760 -----FPV-PLSQCP-----MSLPTTSPPTSP-CSIQVLS 790
 1198 STALYPSVYATSASAGSSTATPGKPPAVASQQAAGSTVGTALTSVSTTSPSTAS 1257
 791 SPFP-----PCC-----PETSFPSTTA 808
 1258 QLSIQSSSTSTPLTARTVVSASLDTKSHSSTTGAFSLASSSSSPGAGVSYISQ 1317
 809 APILSLASAFSL--AVMTVAQSL-----LAFS-----TSSSSPG----- 840
 1318 PGGLHPIVTSVLAETPLPQAAAPTSTPLLPQVPSIPLVQPVAVYAVQOQLIHQOPQ 1377
 841 -----TLSP-----GNPSPGPIBPPIPTTSPCC-----HSPS 873
 1378 PALLPNQPHHCPEVSDTQPKAPGIDIKLEBKLSLSEHSSGAQHASVLETSLV 1437
 874 P-----FSPISQVSNPSPHPTSP 895
 1438 IESTVTEGIP-PAVASKLITSTSCLEPTNLPLCTVALPVPVTPGVSTFVTTT 1426
 896 PFSSTSEFPVPLSQCPWSSILPTTSPPTSPQVGTAVSTYPPSGKSSAAAPKWTGA 955
 1497 SGVPEGT-----APSKPPLTKAPVLPGVTELPAGTLPSOGLPPFPSPSTLOS 1543
 956 GQRPRLHLLLELLELLEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEHLEH 1009
 1544 QQPLEDIDAQRLTSLSEKLTTSVAVGVSMAAPATATEAGTORQKGVSGVKEGVLIATS 1603
 1010 PQ-----IPLHPPAL----- 1020
 1604 SGAGVFCKGRFQVSVADGAQKKGKNSBDKSVHFPSSSTSESSVLSSSPESTLVKPER 1663
 1021 -----HEHSPPAHDSFRSHS----- 1035
 1664 NGITIPGISDVESAHKTTASEAKSDTQPTKVGARQVTTTANKVGRFVSKTEDKITD 1723
 1036 --LSVPGVLS--FRLHLRL-----VGRQVTSKBPAPPLLOPTSPSTLSG 1079
 1724 TKKRGVAVSPFMDLEQAVIPAVIPKKEKPELSEPHLNGPSSDPEAFLSRVDVDSGS 1783
 1080 S-----PKSPPTLT-----SESDTE-----DSAGGG 1102
 1784 PHSPHQLSSKSLPSQNLSSLSNSFNSSVMSNEDIDEDLKLRLRDLKHLEKIQD 1843
 1103 P-----ERHLSVET 1112
 1844 LOSROKHELESYTKLGKVPVAVIIPPAALSGRRRRPTKSGS-KSSRSSSLGKNSPOL 1902
 1113 LQTLQKKEIEDLXSRGKQPPGIVAPAAMLSSQR--LSKGSFTRSRNSQIORSFPC 1170
 1903 SGNLSSGASAVLHPQOQLHPGNIPESQNL-----LQPLKPSSSDNTYSAFTSD 1955
 1171 VPAIPATQEAFA-----GEFLEPGRRLQNSPSALRONSSSPSHNLFITSGK 1220
 1956 GAISVPSLSAPGQSTSTNVGATVVSQAQAQAPPMNT 1993
 1221 -----PLTSATSRWMSFHT--SLGSMFILTCKPMYT 1250

```

RESULT 15
US-10-415-011-17
; Sequence 17, Application US/10415011
; Publication No. US2004005394A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: XU, Yuning
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: DING, Li
; APPLICANT: TANG, Y. Tom
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LU, Yan
; APPLICANT: YUE, Henry
; APPLICANT: BURFORD, Neil
; APPLICANT: BANDMAN, Olga
; APPLICANT: TRIBOUTY, Catherine M.
; APPLICANT: TAL, Preeti G.
; APPLICANT: RECIBON, Shirley A.
; APPLICANT: LU, Dyung Anna M.
; APPLICANT: BOROMSKY, Mark L.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: SWARNAKER Anita
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: KHAN, Farrah A.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0262 USN
; CURRENT APPLICATION NUMBER: US/10/415, 011
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: PCT/US01/47728
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/242, 410
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/244, 068
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/245, 708
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/247, 672
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249, 565
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252, 730
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250, 807
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ. ID NOS: 44
; SOFTWARE: PERL Program
; SEQ. ID NO 17
; LENGTH: 1084
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US2004005394A1 55009053CD1
US-10-415-011-17

```

Query Match 18.94; Score 2038.5; DB 12; Length 1084;

Best Local Similarity 32.84; Pred. No. 1,1e-93; Mismatches 312; Indels 673; Gaps 50;

Matches 566; Conservative 172; Mismatches 312; Indels 673; Gaps 50;

207 LETKAVMSNDGFLKEDIEIGSGFKTVYKGLDTETTVAVANCELODRKLTSESRQRFK 266

1 METQAVATSPDGRKTKDIDIGSGFKTVYKGLDTETTVAVANCELODRKLTSESRQRFK 60

267 EBAEMKGLQHPNIVRFYDSWESTVKGKCIIVLVTIMTSGTLKTYIKRPFVKIKIVLRS 326

```

Db 61 EEVEMLKGLQHPNIVRFYDSWESTVKGKCIIVLVTIMTSGTLKTYIKRPFVKIKIVLRS 120
Qy 327 WCKOILKGLQHPNIVRFYDSWESTVKGKCIIVLVTIMTSGTLKTYIKRPFVKIKIVLRS 386
Db 121 WCKOILKGLQHPNIVRFYDSWESTVKGKCIIVLVTIMTSGTLKTYIKRPFVKIKIVLRS 180
Qy 387 PEMAPMEYEEKYDESDVYAFQCMLEMAISEPYSECONAQAIIYRYVTSQVPAFDK 446
Db 181 PEMAPMEYEEKYDESDVYAFQCMLEMAISEPYSECONAQAIIYRYVTSQVPAFDK 240
Qy 447 VAIPVEKEITIEGTRONKDERYSIKDLNHAFFOEETGVYVLEAEEDDGEKIAIKMLRT 506
Db 241 VKIPEVKEITIEGTRONKDERYSIKDLNHAFFOEETGVYVLEAEEDDGEKIAIKMLRT 300
Qy 507 EDIKKLKGYKQNEAIEFSFDLERDVPEDVQOMVSGYVEGDHHTMAKAIIDRVSLK 546
Db 301 EDARR-GKPRDQALIEFLFQGRDAEYVQOMVGLGVCEADYQVPAVAVERYVAAIQ 359
Qy 567 RKRECRQVREBEQEKKESSILKQOVESASQTIKOLPSASTGIPYASTTSASVSTQ 626
Db 360 RKREKLRKARE-----LEALP-PEGPPATVPMAGPSP 393
Qy 627 V--BPPEPADQHQLOQYQPSI-SVLSGTVYDSQSGSVFTSESVSSQQTIVSYSGQH 682
Db 394 VPPEPEPEPADQHQPFRRHASYSTTDCETDGYLSSGPFADADPALQ----- 445
Qy 683 QAHSTGVPHPTSTYQASQPHGVYPPSSVAQGSQGPSSSLTGVSSSQPIQHPOQ 742
Db 446 -----DGVVPSL--AESHLCLPSA----- 463
Qy 743 QGIQTAPEQQTVOYSLQSTSTSEATTAQPVSPQAPQVLPQVSAKQSTQGVSAVAP 802
Db 464 -----FALS-----IPSGGSPSPSP----- 480
Qy 803 EPVAVAPQATQPTTIASSVDASHSPVAGMSDGNENVPSSSGRHEGRTTKRHYKSVRS 862
Db 481 -----DSYASDAASGLSVGEGM-GQMRPPGGRNLR-----RP 513
Qy 863 RSRHEKTSRPKLRLIVSNKGRVYVCELETHNRKRVTRKEDLDGNPEELATIMYNNDF 922
Db 514 RSR-----LVTVSVSDNDVVECOQTHNSKRVTRFPEDLDGSGPEELIAAMAYNEF 565
Qy 923 IIAIERESEVDQVREIIEKXADEMLSEDVSEPEGDGLESLOCKDDYGFSGSKLEGEFK 982
Db 566 ILPSEHDGFLRRREILQVETILKRDGTGMEAEADTLSPQ----- 607
Qy 983 QPIPASMFQOIGITPTSSILQVYVHAGRFIVSPVESKLRBSKVPPSEITDTVAASTAQ 1042
Db 608 EPAFLPALVPPPLDPENBEIQ-----SSTSLERHMTASTSSSGGTPL 652
Qy 1043 SPGMNLSHASLSLQAFSELRAQMTGPNTPAPNFSGHTGTFVVPFLSSIGVPT 1102
Db 653 SPG-NPFSPEPTPISSPEPIF-----PITSPP-CHPSPS--PSPDISQVSSNRS 686
Qy 1103 TAAATAPVPATSPSPNDISTSVYQSEVTVPTBEGIAGVATSTGVTSGLPIPPVSESPV 1162
Db 697 PHTSSPLPSSSTPE-----PSTSPSTTAAPLUSLAS-----FV-PLSQCP- 721
Qy 1163 LSSVSSITTPAVVSGITSPSLQVPTSTSEIVSVSTALVPSVTSATASAGASTATPG 1222
Db 722 -----WSSLPTTSPPTSPSP-CSIQVTLSSPFF----- 747
Qy 1223 PKPFAVSSQQAAGSTTVGATLTSTSTSPSPASQSLQSSSTSTPTLATVTVVSAHS 1282
Db 748 --PFC-----PSTSPSTTAAPLUSLAS----- 769
Qy 1283 LDKTSHSSTTGLAFSLASABSSSSPGAGVSVYSQPGGLHPIVPSVIASTPILPOAAGP 1342
Db 770 -----AFSLA----- 774
Qy 1343 TSTPLLPQVPSITPPLVQVPAVYVQOTLIHSQPOLLPNQHTHCPEVSDSTQKAPG 1402

```

```
Db 775 -----VMTAOSL--SPSPGLSSGP----- 794
QY 1403 IDDKTEELKRLSFSEHSSGAQHASVSLTSLVISTVTGIPGTAVAPSKLLTSTTS 1462
Db 795 -----PA----- 796
QY 1463 TCUPTNLPLGTVALPVTPVTPGVSTPVSSTTSQVKGPGTAPSKPPLTKAPVLPVGTSL 1522
Db 797 ---PSPLP---SLPLPPVAPGGQSP--SPHTAEVB---SEASPP----- 832
QY 1523 PAGTLPEQLPPFPQSLTQSQPLEDLDAQLRRTLSPEXITVTSAVGPVSMAPTAITE 1582
Db 833 PARLPGE-----ARLAPIS----- 847
QY 1583 AGTQPOKVGQVKEGCVLATSSGAGVPMGRFQVSVAAADGAQKEGKNSKSEDAKSVHFSS 1642
Db 848 -----EEGKPOL-----VGRFOVT-----SSKEPAEPLPLQPT 875
QY 1643 TSASSVLSSSPPESTLVKPEPMGITIPGISDVPESA---HKTASEAKSDTGQPTKVR 1699
Db 876 ---SPTLSGS-----PKPSTPQLTSSSDTEDSAGGGETREALAESD----- 915
QY 1700 FQYTTANKVGRFSVSKTEKIDTKKEGVPVAPPPMDLEQAVLPAVTPKKEKPELSEPS 1759
Db 916 ---RAABGLG-----AGVEEEG----- 929
QY 1760 HUNGSSSDPEAAFLSRVDVDDGSGSPH---SPHOLSSKLPSONLSQSLSNSFNSSYMSD 1816
Db 930 -----DDGK-EPQVGSFPQPLSHSPVWMNYSYS-----SLCLSEB 965
QY 1817 NESDIEDDEDLKLEIRLRDLRHLKEIQDQROKHEIESLYTKLGKVPVAVIIPPAAPLSG 1876
Db 966 SESSGEDEEFMAELQSLRQKHLSVEVTLQTLQKKEIEDLYSRIGKQPPGIVAPAAMLSS 1025
QY 1877 RRRRPTSKGS-KSSRSSLGNKSPQLSG-----NLSGQSAAS 1913
Db 1026 RQRR--LSKGSFPTSRNSLQRESEPPGPGIMRRNSLSSSTGS 1066
```

Search completed: September 24, 2004, 01:25:47
Job time : 232 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

Run on: September 24, 2004, 01:03:14 ; Search time 41 Seconds
(without alignments)
2689.587 Million cell updates/sec

Title: US-10-010-720-14
Perfect score: 10812
Sequence: 1 MSGGNAEKGSSTPGSLFLSP.....NISMLQKSIISNPPGSMURTT 2136

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : Issued Patents AA:*
1 : /cgn2.6/prodata/2/aa/5A_COMB.pep:*
2 : /cgn2.6/prodata/2/aa/5F_COMB.pep:*
3 : /cgn2.6/prodata/2/aa/5A_5B_COMB.pep:*
4 : /cgn2.6/prodata/2/aa/6F_COMB.pep:*
5 : /cgn2.6/prodata/2/aa/PTCTUS_COMB.pep:*
6 : /cgn2.6/prodata/2/aa/backfile1.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	10810	100.0	2136	4	US-09-854-856-14	Sequence 14, App1	
	2	210753.5	99.5	2229	4	US-09-854-856-8	Sequence 8, App1	
	3	10677	98.8	2382	4	US-09-854-856-2	Sequence 2, App1	
	4	10646	98.5	2108	4	US-09-854-856-30	Sequence 30, App1	
	5	10589.5	97.9	2201	4	US-09-854-856-24	Sequence 24, App1	
	6	10513	97.2	2354	4	US-09-854-856-18	Sequence 18, App1	
	7	10508	97.2	2076	4	US-09-854-856-16	Sequence 46, App1	
	8	10451.5	96.7	2169	4	US-09-854-856-40	Sequence 40, App1	
	9	10375	96.0	2322	4	US-09-854-856-34	Sequence 34, App1	
	10	10344	95.7	2048	4	US-09-854-856-62	Sequence 62, App1	
	11	10287.5	95.1	2141	4	US-09-854-856-56	Sequence 56, App1	
	12	10211	94.4	2294	4	US-09-854-856-50	Sequence 50, App1	
	13	9929	91.8	1999	4	US-09-854-856-16	Sequence 16, App1	
	14	9872.5	91.3	2092	4	US-09-854-856-10	Sequence 10, App1	
	15	9796	90.6	2245	4	US-09-854-856-4	Sequence 4, App1	
	16	9765	90.3	1971	4	US-09-854-856-32	Sequence 32, App1	
	17	9708.5	89.8	2064	4	US-09-854-856-26	Sequence 26, App1	
	18	9632	89.1	2217	4	US-09-854-856-20	Sequence 20, App1	
	19	9627	89.0	1939	4	US-09-854-856-48	Sequence 48, App1	
	20	9570.5	88.5	2032	4	US-09-854-856-42	Sequence 42, App1	
	21	9494	87.8	2185	4	US-09-854-856-36	Sequence 36, App1	
	22	9463	87.5	1911	4	US-09-854-856-64	Sequence 64, App1	
	23	9406.5	87.0	2004	4	US-09-854-856-58	Sequence 58, App1	
	24	9330	86.3	2157	4	US-09-854-856-52	Sequence 52, App1	
	25	4053	37.5	982	4	US-09-854-856-6	Sequence 6, App1	
	26	4043	37.4	829	4	US-09-854-856-12	Sequence 12, App1	
	27	3889	35.0	954	4	US-09-854-856-22	Sequence 22, App1	

28	3879	35.9	801	4	US-09-854-856-38	Sequence 28, Appl
29	3751	34.7	922	4	US-09-854-856-38	Sequence 28, Appl
30	3741	34.6	769	4	US-09-854-856-34	Sequence 44, Appl
31	3587	33.2	894	4	US-09-854-856-44	Sequence 54, Appl
32	3577	33.1	741	4	US-09-854-856-60	Sequence 60, Appl
33	1677	15.5	324	4	US-09-205-258-33	Sequence 53, Appl
34	1626	15.0	1601	4	US-09-344-473E-40	Sequence 40, Appl
35	1323	12.2	309	4	US-09-344-473E-37	Sequence 37, Appl
36	1280.5	11.8	1381	4	US-09-808-701A-25	Sequence 25, Appl
37	761.5	7.0	613	4	US-09-344-473E-39	Sequence 39, Appl
38	745.5	6.9	516	4	US-09-344-473E-34	Sequence 34, Appl
39	702.5	6.5	549	4	US-09-344-473E-32	Sequence 32, Appl
40	678.5	6.3	677	4	US-09-344-473E-38	Sequence 38, Appl
41	672	6.2	133	4	US-09-205-258-534	Sequence 534, Appl
42	649	6.0	138	1	US-08-664-596B-22	Sequence 22, Appl
43	439	4.1	2468	4	US-09-976-1594-726	Sequence 726, Appl
44	411	3.8	1259	4	US-09-688-188B-13	Sequence 13, Appl
45	411	3.8	1259	4	US-09-291-417D-13	Sequence 13, Appl

ALIGNMENTS

```

RESULT 1
US-09-854-856-14
: Sequence 14, Application US/09854856
: Patent No. 6541252
: GENERAL INFORMATION:
: APPLICANT: Walke, D. Wade
: APPLICANT: Hilbun, Erin
: APPLICANT: Donoho, Gregory
: APPLICANT: Turner, C. Alexander Jr.
: TITLE OF INVENTION: No. 6541252e1 Human Kinases and Polynucleotides
: TITLE OF INVENTION: Encoding the Same
: FILE REFERENCE: LEX-0178-USA
: CURRENT APPLICATION NUMBER: US/09/854,856
: CURRENT FILING DATE: 2001-05-14
: PRIOR APPLICATION NUMBER: US 60/206,015
: PRIOR FILING DATE: 2000-05-19
: NUMBER OF SEQ ID NOS: 64
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14
: LENGTH: 2136
: TYPE: PR1
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)...(2136)
: OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-14

Query Match      100.0%; Score 10810; DB 4; Length 2136;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSGAAEKOSSTPSSLFLSPAPAPAKNGSSSDSSVGEKLGAAAADAVTGRTEYRRRRRT 60
Db      1 MSGGAEEKOSSTPSSLFLSPAPAPAKNGSSSDSSVGEKLGAAAADAVTGRTEYRRRRRT 60

QY      61 MDKDSRGAAATTTTTHRRFFRRSYICDSNATALELPGLPJSLPQPSIPAAVPOQAPPEPH 120
Db      61 MDKDSRGAAATTTTTHRRFFRRSYICDSNATALELPGLPJSLPQPSIPAAVPOQAPPEPH 120

QY      121 REEYVTAATATSOVAAQPPAAAPGQAVAGAPSTVPSSTSKDRPVSOPLVGSKEEPPP 180
Db      121 REEYVTAATATSOVAAQPPAAAPGQAVAGAPSTVPSSTSKDRPVSOPLVGSKEEPPP 180

QY      121 REEYVTAATATSOVAAQPPAAAPGQAVAGAPSTVPSSTSKDRPVSOPLVGSKEEPPP 180
Db      121 REEYVTAATATSOVAAQPPAAAPGQAVAGAPSTVPSSTSKDRPVSOPLVGSKEEPPP 180

QY      181 ARSGGGGSAKEPOERSQOQDDIELEETKAVGMSNDRFLKFDIEIGRGSFKTYKGLD 240
Db      181 ARSGGGGSAKEPOERSQOQDDIELEETKAVGMSNDRFLKFDIEIGRGSFKTYKGLD 240

QY      241 TETTYEAVWCELODRKLTLSRQRFKEAEMLKGLQHNNITRFQDSWESTYKGGKCIIVL 300
Db      241 TETTYEAVWCELODRKLTLSRQRFKEAEMLKGLQHNNITRFQDSWESTYKGGKCIIVL 300

```

Db 241 TETVEVAMCEIQQDKLTKSERQFKEBAEMLKGLQHPNIVRFYDWSWETVKKCIYLV 300
 QY 301 TELMTSSGLKTYLTKRFKXMKTKVLBSNGROIILKGLQFHTRPPIIHRDLKCDNIPTNGP 360
 Db 301 TELMTSSGLKTYLTKRFKXMKTKVLBSNGROIILKGLQFHTRPPIIHRDLKCDNIPTNGP 360
 QY 361 TGSVXIGLGLATLKRASFVSVIGTPEPFMAPEMEYEKYDESVDVYAFAGMCMLEAVATSEY 420
 Db 361 TGSVXIGLGLATLKRASFVSVIGTPEPFMAPEMEYEKYDESVDVYAFAGMCMLEAVATSEY 420
 QY 421 PYSEQONAAQIYRRVTSVGVKPAFDPKVAIPEYKEIIIEGCIROMKDERYSIKDLNHAFFQ 480
 Db 421 PYSEQONAAQIYRRVTSVGVKPAFDPKVAIPEYKEIIIEGCIROMKDERYSIKDLNHAFFQ 480
 QY 481 EETGVRAELAEEDDOEKIAIKMLRIEDIKLKGKXONKAEIPEFDERVPEVNAQEM 540
 Db 481 EETGVRAELAEEDDOEKIAIKMLRIEDIKLKGKXONKAEIPEFDERVPEVNAQEM 540
 QY 541 VESGVCEGDHKTAKAIKDRVSLIKRKEQROLVREOEKKKQESSLKOQVEQSSASQ 600
 Db 541 VESGVCEGDHKTAKAIKDRVSLIKRKEQROLVREOEKKKQESSLKOQVEQSSASQ 600
 QY 601 TGTIQLRBASTGCTIPLASTTSASVSTQVPEPEADQHQOLQYQPSISVLSGTVDSGQ 660
 Db 601 TGTIQLRBASTGCTIPLASTTSASVSTQVPEPEADQHQOLQYQPSISVLSGTVDSGQ 660
 QY 661 SSVFTESVSSQOQVSVGSQHEOAHSTGTVPNGHPTTQAOQSPHGVYPPSSVAAQSGSQ 720
 Db 661 SSVFTESVSSQOQVSVGSQHEOAHSTGTVPNGHPTTQAOQSPHGVYPPSSVAAQSGSQ 720
 QY 721 QPSSSLTGVSSSQPIQHPOQOOGIOQTAPPOQTQVXSLSQTSSEATTPAQVPOQAP 780
 Db 721 QPSSSLTGVSSSQPIQHPOQOOGIOQTAPPOQTQVXSLSQTSSEATTPAQVPOQAP 780
 QY 781 QVLPQVAGAKSGTQGVSOVAPAEFVAVAOPOATPPTIASSVDASHSDVAGMSDGENV 840
 Db 781 QVLPQVAGAKSGTQGVSOVAPAEFVAVAOPOATPPTIASSVDASHSDVAGMSDGENV 840
 QY 841 PSSSGRHEGRTTKHYRKSVAERSRHEKTSRPKRIINVSNGRVRVECOLETNRKMYT 900
 Db 841 PSSSGRHEGRTTKHYRKSVAERSRHEKTSRPKRIINVSNGRVRVECOLETNRKMYT 900
 QY 901 FKFDLDGDNPEELATIMVNDPILAIERESFVDQVREIIEKADMLSEDSVPEEGQGL 960
 Db 901 FKFDLDGDNPEELATIMVNDPILAIERESFVDQVREIIEKADMLSEDSVPEEGQGL 960
 QY 961 ESLOGKDDYGFSGQKLEGEFKQPIPASMPQOIGITPSSITQVHAGRRFIVSPVES 1020
 Db 961 ESLOGKDDYGFSGQKLEGEFKQPIPASMPQOIGITPSSITQVHAGRRFIVSPVES 1020
 QY 1021 RLRESKTFPSEITDTTVAASTAOSFGMNLSSHASLSLQOAFSELRRKOMTEGPTAPNF 1080
 Db 1021 RLRESKTFPSEITDTTVAASTAOSFGMNLSSHASLSLQOAFSELRRKOMTEGPTAPNF 1080
 QY 1081 SHTGPTFPVPPPLSSIAGVPTTAATAPVATSSPNDISTSVYQSEVTPTEGJAGV 1140
 Db 1081 SHTGPTFPVPPPLSSIAGVPTTAATAPVATSSPNDISTSVYQSEVTPTEGJAGV 1140
 QY 1141 ATSGVATSGGLPPIPPSESPVLSVVSITTPAVNSITTSBGLQVPTSTSEIIVASTA 1200
 Db 1141 ATSGVATSGGLPPIPPSESPVLSVVSITTPAVNSITTSBGLQVPTSTSEIIVASTA 1200
 QY 1201 LYSVTVSATSASAGSTATPGPKPVPVSOQAAGSTTVGATLTVSVTTSPSTASQLS 1260
 Db 1201 LYSVTVSATSASAGSTATPGPKPVPVSOQAAGSTTVGATLTVSVTTSPSTASQLS 1260
 QY 1261 IQSSSSTSTPTLAETVAVASHSLDKTSSSTTGLAFSLAPSSSSSPGAGVSSYISQPG 1320
 Db 1261 IQSSSSTSTPTLAETVAVASHSLDKTSSSTTGLAFSLAPSSSSSPGAGVSSYISQPG 1320
 QY 1321 LHLPLVPSVASTPILPOAAGPTSTPLLDVOPSPILPVQPVANVPAVOQTLHSQOPAL 1380
 Db 1321 LHLPLVPSVASTPILPOAAGPTSTPLLDVOPSPILPVQPVANVPAVOQTLHSQOPAL 1380

QY 1381 LPNOPHTHCEPVSDTOPKAPGIDDIKTLKEKLSLSESHSSSGAQAHSVLETSLVIES 1440
 Db 1381 LPNOPHTHCEPVSDTOPKAPGIDDIKTLKEKLSLSESHSSSGAQAHSVLETSLVIES 1440
 QY 1441 TVTGPITTAAPSKILSTSTGCLPPTNLPLGVAALPVTVPVPGOVSTVSTTSYVK 1500
 Db 1441 TVTGPITTAAPSKILSTSTGCLPPTNLPLGVAALPVTVPVPGOVSTVSTTSYVK 1500
 QY 1501 PGTAPSKPPLTKAVLPVTELPAGTLPSEQLPPEPGPSILTOQOPLDLDLAOLRTTSP 1560
 Db 1501 PGTAPSKPPLTKAVLPVTELPAGTLPSEQLPPEPGPSILTOQOPLDLDLAOLRTTSP 1560
 QY 1561 EXITVTGAVGVSAAPALITAEAGTOPQKGVQVYKEGVALTSSGAGYFKRGPVSVAA 1620
 Db 1561 EXITVTGAVGVSAAPALITAEAGTOPQKGVQVYKEGVALTSSGAGYFKRGPVSVAA 1620
 QY 1621 DGAQKEGNKSEDAKSVFSESTSESSVLSSSSPSTLVKPEPNGITTPGISDVPESAH 1680
 Db 1621 DGAQKEGNKSEDAKSVFSESTSESSVLSSSSPSTLVKPEPNGITTPGISDVPESAH 1680
 QY 1681 KTTASBAKSDTQPTKYGRFQVTTTANKVGRFVSXKTEDKITDTKKEGVPVAPPEFMDLEQ 1740
 Db 1681 KTTASBAKSDTQPTKYGRFQVTTTANKVGRFVSXKTEDKITDTKKEGVPVAPPEFMDLEQ 1740
 QY 1741 AVLPAAVLPKKEKPELSEPSHLNGPSSDPEAFLSRDVEDGSGSPHSPHQLSSKSLPSQNL 1800
 Db 1741 AVLPAAVLPKKEKPELSEPSHLNGPSSDPEAFLSRDVEDGSGSPHSPHQLSSKSLPSQNL 1800
 QY 1801 SOSLSNFSNYSNDSNEDSIEDLKLRLRLBDKHLKEIODLSQKHEIESLYTKLG 1860
 Db 1801 SOSLSNFSNYSNDSNEDSIEDLKLRLRLBDKHLKEIODLSQKHEIESLYTKLG 1860
 QY 1861 KVPAAVILPPAALPSGRRRPPTKSGKSSRSSLGANKSPOLSGNLSGOSAAVYLPQOT 1920
 Db 1861 KVPAAVILPPAALPSGRRRPPTKSGKSSRSSLGANKSPOLSGNLSGOSAAVYLPQOT 1920
 QY 1921 LHPGNIPESGQOOLQPLKPSPSDMLYSAFTSDGAISSPISAPQOQTSSTTVGATV 1980
 Db 1921 LHPGNIPESGQOOLQPLKPSPSDMLYSAFTSDGAISSPISAPQOQTSSTTVGATV 1980
 QY 1981 NSQAQAQAPAMTSSRGFTDLDLKLVDNWARDAMLSGRGSGKGMNTEGPMARKFS 2040
 Db 1981 NSQAQAQAPAMTSSRGFTDLDLKLVDNWARDAMLSGRGSGKGMNTEGPMARKFS 2040
 QY 2041 APQOLCISMTSNLGGAPISAAATSLGHFTKSMCPQOYGFPATPPGAQWSTGGBAPQ 2100
 Db 2041 APQOLCISMTSNLGGAPISAAATSLGHFTKSMCPQOYGFPATPPGAQWSTGGBAPQ 2100
 QY 2101 PLGQPOVGTASLONFISNLOKISINPPGSNRTT 2136
 Db 2101 PLGQPOVGTASLONFISNLOKISINPPGSNRTT 2136

RESULT 2
 US-09-854-856-8
 ; Sequence 8, Application US/09854856
 ; Patent No. 6541252
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hilbun, Erin
 ; APPLICANT: Donoho, Gregory
 ; APPLICANT: Turner, C. Alexander Jr.
 ; TITLE OF INVENTION: No. 6541252 Human Kinases and Polynucleotides
 ; TITLE OF INVENTION: Encoding the Same
 ; FILE REFERENCE: LEX-0178-USA
 ; CURRENT APPLICATION NUMBER: US/09/854,856
 ; PRIOR FILING DATE: 2001-05-14
 ; PRIOR APPLICATION NUMBER: US 60/206,015
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8


```

; LENGTH: 2229
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2229)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-8

```

```

Query Match      99.5%; Score 10753.5; DB 4; Length 2229;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

```

```

QY 1 MSGGAAGKOSSTPSLFLSPAPAPAKNGSSSDSVGKLGAAADATVGRTEYRRRRHT 60
Db 1 MSGGAAGKOSSTPSLFLSPAPAPAKNGSSSDSVGKLGAAADATVGRTEYRRRRHT 60
QY 61 MDKDSRGAATTTTTHRRFRRSVICSNAATALEPLGLPLSLPQPSIPAAVPGAPPEPH 120
Db 61 MDKDSRGAATTTTTHRRFRRSVICSNAATALEPLGLPLSLPQPSIPAAVPGAPPEPH 120
QY 121 REETVATATSOVAQOQPPAAAPEQAVAGPAPSTVSSSKDPVPQPSLVGSKBPP 180
Db 121 REETVATATSOVAQOQPPAAAPEQAVAGPAPSTVSSSKDPVPQPSLVGSKBPP 180
QY 181 ARSGGGGSAKEPQEESSQODDIEELTKAVGMSNDGRFLKFDIEIGRGSFKTVYKGLD 240
Db 181 ARSGGGGSAKEPQEESSQODDIEELTKAVGMSNDGRFLKFDIEIGRGSFKTVYKGLD 240
QY 241 TETTVAVAMCELQDRKLTSEORFKBEAEMLKGLQHPNIVRFYDSWSTVKGKCIYLV 300
Db 241 TETTVAVAMCELQDRKLTSEORFKBEAEMLKGLQHPNIVRFYDSWSTVKGKCIYLV 300
QY 301 TELMTSGTLKTYLKRPFYMKIKVIRSMCRQLKGLQPLHTPTPIHRDLKCDNIPTGP 360
Db 301 TELMTSGTLKTYLKRPFYMKIKVIRSMCRQLKGLQPLHTPTPIHRDLKCDNIPTGP 360
QY 361 TGSYKIGLGLATIKRASFAKSVIGTFPMAPEMEYEEKYDSVVYVAFGCMLEMAISEY 420
Db 361 TGSYKIGLGLATIKRASFAKSVIGTFPMAPEMEYEEKYDSVVYVAFGCMLEMAISEY 420
QY 421 PYSECQNAQAQYRRVTSVGPASFDKVAIPVKEIIIEGCIROKNDERYSIKDLINHAFFQ 480
Db 421 PYSECQNAQAQYRRVTSVGPASFDKVAIPVKEIIIEGCIROKNDERYSIKDLINHAFFQ 480
QY 481 BETGVAVLAEDDEGEKTAIKMLRIEDIKLKGGYKKNELIERSFDLERVPEDEVAGEM 540
Db 481 BETGVAVLAEDDEGEKTAIKMLRIEDIKLKGGYKKNELIERSFDLERVPEDEVAGEM 540
QY 541 VESGVCEGDHKTMAKAIKDRVSLIKRREORQVREBQEKKESSSLKQOVESSASQ 600
Db 541 VESGVCEGDHKTMAKAIKDRVSLIKRREORQVREBQEKKESSSLKQOVESSASQ 600
QY 601 TGIKQLPSASTGIPATSTTSASVSTQVEPEEADQHOQLQYQPSISVLSDGTVDGSGG 660
Db 601 TGIKQLPSASTGIPATSTTSASVSTQVEPEEADQHOQLQYQPSISVLSDGTVDGSGG 660
QY 661 SSVFRESKVSQQTYSYSGHQAHSVTGVGHIPSTVQAOSQPHGVPPSSVAGQSGQG 720
Db 661 SSVFRESKVSQQTYSYSGHQAHSVTGVGHIPSTVQAOSQPHGVPPSSVAGQSGQG 720
QY 721 QPSSSLTGVSSSQIPIHQPOQOQIQOTAPPOQTQVYSICTSTSEATTAQPVQAPAP 780
Db 721 QPSSSLTGVSSSQIPIHQPOQOQIQOTAPPOQTQVYSICTSTSEATTAQPVQAPAP 780
QY 781 QVLPQVSAAGKQ----- 791
Db 781 QVLPQVSAAGKQ----- 791
QY 791 QVLPQVSAAGKQ----- 791
Db 791 QVLPQVSAAGKQ----- 791
QY 841 VVQPVESNLIVPMGVGGQVQVSGGSLAQAPTTSSQAVLESITQGVSAVAPAEVAV 900
Db 841 VVQPVESNLIVPMGVGGQVQVSGGSLAQAPTTSSQAVLESITQGVSAVAPAEVAV 900

```

```

QY 808 ACPQATOPTTLASSVDASDVASGMSDGNENPVSSSGRRERGTTRKHYKSVRSRHE 867
Db 901 ACPQATOPTTLASSVDASDVASGMSDGNENPVSSSGRRERGTTRKHYKSVRSRHE 960
QY 868 KTSRPKRLITNVNKNKDRVVECOLETHNRKMTFKFDLODNDPEEATIANVNDPILAE 927
Db 961 KTSRPKRLITNVNKNKDRVVECOLETHNRKMTFKFDLODNDPEEATIANVNDPILAE 1020
QY 928 RESFVQVREIIEKADMLSEDSVEPEGQGLSEIQGKDQYFGSGOKLEGEKQIPPA 987
Db 1021 RESFVQVREIIEKADMLSEDSVEPEGQGLSEIQGKDQYFGSGOKLEGEKQIPPA 1080
QY 988 GSMPOQIGIPTSSLTQVHSAGRFIVSPVESRLBRSKVPESBITDTVAASTAQSQGM 1047
Db 1081 GSMPOQIGIPTSSLTQVHSAGRFIVSPVESRLBRSKVPESBITDTVAASTAQSQGM 1140
QY 1048 LSHSASLSLQOAFSELRRAQMTGPTAPPNESHGTPFPVPPFLSIAGVTTAAAT 1107
Db 1141 LSHSASLSLQOAFSELRRAQMTGPTAPPNESHGTPFPVPPFLSIAGVTTAAAT 1200
QY 1108 APVPATSPNDISTSVIQSEVTPTEEGIAGVATSTGVVTSGLPIPVSESPVLSVV 1167
Db 1201 APVPATSPNDISTSVIQSEVTPTEEGIAGVATSTGVVTSGLPIPVSESPVLSVV 1260
QY 1168 SSITTPAVASISTTSPSLQVPTSTSEIVSSTALYPSEVTVSATASAGGSTATPGPRPA 1227
Db 1261 SSITTPAVASISTTSPSLQVPTSTSEIVSSTALYPSEVTVSATASAGGSTATPGPRPA 1320
QY 1228 VVSGQAAGSTTGAATLTVSSTTTSPSTASOLSTQSSSTPTPLAFVWVSAHSLKTS 1287
Db 1321 VVSGQAAGSTTGAATLTVSSTTTSPSTASOLSTQSSSTPTPLAFVWVSAHSLKTS 1380
QY 1288 HSTTGLAFSLAPSSSSPGAGVSYISOPGHLPIVTSVIASTPILPQAAGTSTPL 1347
Db 1381 HSTTGLAFSLAPSSSSPGAGVSYISOPGHLPIVTSVIASTPILPQAAGTSTPL 1440
QY 1348 LPQVPSIPPLVQPVANPAVQOQLIHSQPOPALLPNOPHHCEVDSDTQPKAGCIDIX 1407
Db 1441 LPQVPSIPPLVQPVANPAVQOQLIHSQPOPALLPNOPHHCEVDSDTQPKAGCIDIX 1500
QY 1408 TLEKRLSLPSEHSSSGAHSVLETSIVTSGIPPTAAVAPSKLSTSTGCLP 1467
Db 1501 TLEKRLSLPSEHSSSGAHSVLETSIVTSGIPPTAAVAPSKLSTSTGCLP 1560
QY 1468 TNLPLGTVALPVTVPVTPPGQVSTPVSTTSGVKGTAISKXPLTKAPVLPVGTBLPAGTL 1527
Db 1561 TNLPLGTVALPVTVPVTPPGQVSTPVSTTSGVKGTAISKXPLTKAPVLPVGTBLPAGTL 1620
QY 1528 PSEQLPPEPGPSLTQSQOPLLEDLDAQLRRTLSPEXITVTSVAGVSNMAAPTAITEAGQP 1587
Db 1621 PSEQLPPEPGPSLTQSQOPLLEDLDAQLRRTLSPEXITVTSVAGVSNMAAPTAITEAGQP 1680
QY 1588 QKGVSVKEGEPVLAATSSAGVFKMGRFQVSAADGAQKEGKNSKEDAKSVHFESESTSSS 1647
Db 1681 QKGVSVKEGEPVLAATSSAGVFKMGRFQVSAADGAQKEGKNSKEDAKSVHFESESTSSS 1740
QY 1648 VLSSSPESTLVKEPENITIPGISIDVPESAHTTASBAKSDTQOPTKGRFOVTTTAN 1707
Db 1741 VLSSSPESTLVKEPENITIPGISIDVPESAHTTASBAKSDTQOPTKGRFOVTTTAN 1800
QY 1708 KVGRRSVSKTEDEKTDITTKKBPVAPPPMDLEQAVLPAVITPKKKEPELSSESHLNGPSSD 1767
Db 1801 KVGRRSVSKTEDEKTDITTKKBPVAPPPMDLEQAVLPAVITPKKKEPELSSESHLNGPSSD 1860
QY 1768 PEAFLASDVDDGSGSPHSPHOLSKSLPSQVLSOSLSNSFNSSVMSDNESEDIEDDLX 1827
Db 1861 PEAFLASDVDDGSGSPHSPHOLSKSLPSQVLSOSLSNSFNSSVMSDNESEDIEDDLX 1920
QY 1828 LEIARLRKHLKEIQLDQSRQKHETESLYTKLVKVPVAVIIPPAAPLSGRRRRPTKSKGS 1887
Db 1921 LEIARLRKHLKEIQLDQSRQKHETESLYTKLVKVPVAVIIPPAAPLSGRRRRPTKSKGS 1980
QY 1888 KSSRSSSLGKNSPOLSGNLSCQSAASVLAHQOQLHPNANIPESGQNLQPLKPSBSN 1947

```

```

Db      1981 KSSRSSISGKNSPQJSLSGNSGSAASVLPQCTLHPGNIPESSGQNLQPLKPSBSN 2040
QY      1948 LYSATSPGSAISVPSLSAPGQCTSSNTVYGAIVNSQAQAOPPATSSKKGFTTDLHL 2007
Db      2041 LYSATSPGSAISVPSLSAPGQCTSSNTVYGAIVNSQAQAOPPATSSKKGFTTDLHL 2100
QY      2008 VDNMARDAVNLSGRGSKGHNMYEGGMAKFSAPGQCLISMTSNLGGSAPIASAATSL 2067
Db      2101 VDNMARDAVNLSGRGSKGHNMYEGGMAKFSAPGQCLISMTSNLGGSAPIASAATSL 2160
QY      2068 GHFTSMCPPOQYGPATPFGAQMSTGAPAPQIGQFQPVGTASLQNFNLSNLSN 2127
Db      2161 GHFTSMCPPOQYGPATPFGAQMSTGAPAPQIGQFQPVGTASLQNFNLSNLSN 2220
QY      2128 PPGSNLRRTT 2136
Db      2221 PPGSNLRRTT 2229

```

RESULT 3

```

US-09-854-856-2
; Sequence 2, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854, 856
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206, 015
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2382
; TYPE: PRN
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2382)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-2

```

```

Query Match      96.8%; Score 10677; DB 4; Length 2382;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 246; Gaps 1;
QY      1 MSGGAAKKOSTPSTPLSPAPAPKNGSSSDSVGKLGAAADANTGTEERRRRRT 60
Db      1 MSGGAAKKOSTPSTPLSPAPAPKNGSSSDSVGKLGAAADANTGTEERRRRRT 60
QY      61 MDXDSRGAATTTTTEHFRFRRSVICSNAATALELPGCLPLSPPOSIPAAVPOSAPREP 120
Db      61 MDXDSRGAATTTTTEHFRFRRSVICSNAATALELPGCLPLSPPOSIPAAVPOSAPREP 120
QY      121 REETVNTATSOVAQAOPPAAPAEQAVAGPAPSTVSSSTSKORPVSOPLSVGSKKEPP 180
Db      121 REETVNTATSOVAQAOPPAAPAEQAVAGPAPSTVSSSTSKORPVSOPLSVGSKKEPP 180
QY      181 ARSGSGGSAKEPOEERSQOQDDIELETKAVGMSNGRFLKPDIEGRSSFTVYKGLD 240
Db      181 ARSGSGGSAKEPOEERSQOQDDIELETKAVGMSNGRFLKPDIEGRSSFTVYKGLD 240
QY      241 TETTVAVMCELQDKLTSEKRFKEEAMLGLOHPNIVRFYDSWESTVKKKCIIV 300
Db      241 TETTVAVMCELQDKLTSEKRFKEEAMLGLOHPNIVRFYDSWESTVKKKCIIV 300
QY      301 TELMISGLTKYLRKRFVKMKIVLRSWCQILKGLQFLHTRTPPIIHRDLKCNIFITGP 360

```

```

Db      301 TELMISGLTKYLRKRFVKMKIVLRSWCQILKGLQFLHTRTPPIIHRDLKCNIFITGP 360
QY      361 TGSVYKIGLGLATLTKRASFANSVIGTPEFMAPEMYEERYKDESVDYVAGCMLEVAATSEY 420
Db      361 TGSVYKIGLGLATLTKRASFANSVIGTPEFMAPEMYEERYKDESVDYVAGCMLEVAATSEY 420
QY      421 PYSECQNAAQYRRYTSQVKPASFQKVAIPEYKEIIIECIRONKDERYSIKDLNHAFFQ 480
Db      421 PYSECQNAAQYRRYTSQVKPASFQKVAIPEYKEIIIECIRONKDERYSIKDLNHAFFQ 480
QY      481 EETGVAVLAEEDDEKTAIKMLRIEDIKTKGKYKNEALIEFSPDLERDVEDVAQEM 540
Db      481 EETGVAVLAEEDDEKTAIKMLRIEDIKTKGKYKNEALIEFSPDLERDVEDVAQEM 540
QY      541 VESGYVCBGDKRTAKAIKDRVSLIKRREOROLVREEQEKKKQESSLKKQVEOSASQ 600
Db      541 VESGYVCBGDKRTAKAIKDRVSLIKRREOROLVREEQEKKKQESSLKKQVEOSASQ 600
QY      601 TGIKQLPASTGIPASTTSASVSTQVEPEPEADQHQOQYQOPSISVLSGTVDSGQ 660
Db      601 TGIKQLPASTGIPASTTSASVSTQVEPEPEADQHQOQYQOPSISVLSGTVDSGQ 660
QY      661 SSVFTESSVSQQTIVSYGSOHEQAHSTGVPGHISTYQAOSQPHGVPPSSVAQGSQ 720
Db      661 SSVFTESSVSQQTIVSYGSOHEQAHSTGVPGHISTYQAOSQPHGVPPSSVAQGSQ 720
QY      721 QPSSSLTGVSASSQPIQHPQQOQGIQQTAPQQTVOYLSQSTSSSEATTQAQVSPQAP 780
Db      721 QPSSSLTGVSASSQPIQHPQQOQGIQQTAPQQTVOYLSQSTSSSEATTQAQVSPQAP 780
QY      781 QVLPQVSAQKQ----- 791
Db      781 QVLPQVSAQKQPVSPVPPTIQGEPQIPVATQPSVPEVHSGAHLVYGQPLPTPLPQYP 840
QY      792 ----- 791
Db      792 ----- 791
QY      841 VSOQPISTPHVSTAQOTGFSSIPITMAAGIQPLITLASSATTAAIPGVSTVPSQLPTLL 900
Db      792 ----- 791
QY      901 QPVTLQPSQVHPQLQPAVOQSMGIPANLQQAEBVLSGDVLYGFPPLPQYRQDSNI 960
Db      792 ----- 791
QY      961 APSSNVAQVCIHSTVLAPEMPEVLAATPGYFPYVQPYVESNLLVPMGQVGGQVVSQPG 1020
Db      792 ----- 791
QY      1021 GSILAQARTTSSQQAQVLESTQGVQVAPAEVYVAAQOPATQPTTLASSVDSHSDVASGMS 1080
Db      835 DGNENYVSSSGRHHGRTTKKHYKVSRSRHEKTSBPKLIINVSNGDRVVEQOLETH 894
QY      1081 DGNENYVSSSGRHHGRTTKKHYKVSRSRHEKTSBPKLIINVSNGDRVVEQOLETH 1140
Db      895 NRKMTVEKPDLDNDNPELATTIMVNDPILAIRESVVDQVREIIEAADMLSEDSVEP 954
QY      1141 NRKMTVEKPDLDNDNPELATTIMVNDPILAIRESVVDQVREIIEAADMLSEDSVEP 1200
Db      955 EGDQGLSEIQKNDYGRSGSQKLEGEFKQIPASSMPOQIGITPSLTQVYHAGRRFIV 1014
QY      1201 EGDQGLSEIQKNDYGRSGSQKLEGEFKQIPASSMPOQIGITPSLTQVYHAGRRFIV 1260
Db      1015 SPVPESRLRESKYFPESEITDTVAASTAQSFGNMLSHSASSLSQQAFFSELRAAQMTGPN 1074
QY      1261 SPVPESRLRESKYFPESEITDTVAASTAQSFGNMLSHSASSLSQQAFFSELRAAQMTGPN 1320
Db      1075 TAPNESHGPTFPVVPPLFSSITAGVTTAAATAPPAISSPNDISTVYQSEVVPTE 1134
QY      1321 TAPNESHGPTFPVVPPLFSSITAGVTTAAATAPPAISSPNDISTVYQSEVVPTE 1380
Db      1135 EGIAGVATSTGVVTSGLPIPVSESPVLSVVSSTIIPAVVISTTSPLQVPTSTSEI 1194

```

Db 1381 EGIAGVATGTGVTSGGLPIPVSESEFVLSVSSITIPAVVISITTSPSLQVETSTSEI 1440
 QY 1195 VVSSTALVSVTVSATSASAGSSTATPGKPPAVVSOOAGSTTVGATLTSTSTTSPS 1254
 Db 1441 VVSSTALVSVTVSATSASAGSSTATPGKPPAVVSOOAGSTTVGATLTSTSTTSPS 1500
 QY 1255 TAAQLSTIQLSSSTSTPTLAETVVVSAHSLDKTSHSSTTGALFSLAPSSSSSPGAGVSSY 1314
 Db 1501 TAAQLSTIQLSSSTSTPTLAETVVVSAHSLDKTSHSSTTGALFSLAPSSSSSPGAGVSSY 1560
 QY 1315 ISQGGHPLVTVESVIASTPTLLPQAAGPTSTPLLPQVPSIPPLVQPVANPAVQOQLIHS 1374
 Db 1561 ISQGGHPLVTVESVIASTPTLLPQAAGPTSTPLLPQVPSIPPLVQPVANPAVQOQLIHS 1620
 QY 1375 QPOPALPNOPHTCPEVSDTOPKAGIDIDIKLEKRLSEFSEHSSSGAQAHSVSELT 1434
 Db 1621 QPOPALPNOPHTCPEVSDTOPKAGIDIDIKLEKRLSEFSEHSSSGAQAHSVSELT 1680
 QY 1435 SLVIESITVTPGIPPTTAAVAPSKLTSTSTCLPPTNLPGLVALPVPVTPGQVSTPVST 1494
 Db 1681 SLVIESITVTPGIPPTTAAVAPSKLTSTSTCLPPTNLPGLVALPVPVTPGQVSTPVST 1740
 QY 1495 TTSGVKGRTASPKPLTKAPVLPVGTLPAGTLFSEQLPPPPGSLTQSQQLPDLDAQL 1554
 Db 1741 TTSGVKGRTASPKPLTKAPVLPVGTLPAGTLFSEQLPPPPGSLTQSQQLPDLDAQL 1800
 QY 1555 RRTISPEKITVTSVAVGPVMAAPTAITEAGTOPQKGVQVKEGPVLTSSAGVFKMGFR 1614
 Db 1801 RRTISPEKITVTSVAVGPVMAAPTAITEAGTOPQKGVQVKEGPVLTSSAGVFKMGFR 1860
 QY 1615 QVSAVADGAQKEGKNNKSEDAKSVHESSTSESVLSSSPSTLVKEPENGITTPGISSD 1674
 Db 1861 QVSAVADGAQKEGKNNKSEDAKSVHESSTSESVLSSSPSTLVKEPENGITTPGISSD 1920
 QY 1675 VPESAHKTASAEASDTCOPTKVGAFQYTTTANKVGRSVKTEDEKIDTDTKEGPVAPSP 1734
 Db 1921 VPESAHKTASAEASDTCOPTKVGAFQYTTTANKVGRSVKTEDEKIDTDTKEGPVAPSP 1980
 QY 1735 FMDLEQAVLPVAVIPKKEPELSEPSHLLGSPSSDPEAAFLSRDVGDSGPHSPHOLSKS 1794
 Db 1981 FMDLEQAVLPVAVIPKKEPELSEPSHLLGSPSSDPEAAFLSRDVGDSGPHSPHOLSKS 2040
 QY 1795 LPSQNLSQLSNSFNSSYMSSDNESDIEDDKELRLRDKHKEIIDLQSRQKHIES 1854
 Db 2041 LPSQNLSQLSNSFNSSYMSSDNESDIEDDKELRLRDKHKEIIDLQSRQKHIES 2100
 QY 1855 LYTUKLGVPPAVIIPPAVPLSGRRRRPTKSGKSSRSSSGKNSPQLSGNLSCGASAV 1914
 Db 2101 LYTUKLGVPPAVIIPPAVPLSGRRRRPTKSGKSSRSSSGKNSPQLSGNLSCGASAV 2160
 QY 1915 LHPQOTLHPGNIPESGONLLOPKPSPSSDNLVSATSPDAISVPSLSAPGOQTSSTN 1974
 Db 2161 LHPQOTLHPGNIPESGONLLOPKPSPSSDNLVSATSPDAISVPSLSAPGOQTSSTN 2220
 QY 1975 TVGATVNSQAQAQOPPATTSRKGFTTDLHLVDMNARMANLSGRRGSGKHNNYEPG 2034
 Db 2221 TVGATVNSQAQAQOPPATTSRKGFTTDLHLVDMNARMANLSGRRGSGKHNNYEPG 2280
 QY 2035 MARKFSAGQULCISMTSNLGSAPISASATSLGHFTSMCPQOYGPATPFQAWMGST 2094
 Db 2281 MARKFSAGQULCISMTSNLGSAPISASATSLGHFTSMCPQOYGPATPFQAWMGST 2340
 QY 2095 GGAPAPQLGQOPVGTASTLQNFNINSLQKISNPPGSMRLRTT 2136
 Db 2341 GGAPAPQLGQOPVGTASTLQNFNINSLQKISNPPGSMRLRTT 2382

RESULT 4
 US-09-854-856-30
 ; Sequence 30, Application US/09854856
 ; Patient No. 6541252
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade

APPLICANT: Hilbun, Erin
 APPLICANT: Donoho, Gregory
 APPLICANT: Turner, C. Alexander Jr.
 TITLE OF INVENTION: No. 6541252 Human Kinases and Polynucleotides
 TITLE OF INVENTION: Encoding the Same
 FILE REFERENCE: LEX-0178-USA
 CURRENT APPLICATION NUMBER: US/09/854, 856
 PRIOR FILING DATE: 2001-05-14
 PRIOR APPLICATION NUMBER: US 60/206, 015
 NUMBER OF SEQ. ID NOS: 64
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ. ID NO 30
 LENGTH: 2108
 TYPE: PRT
 ORGANISM: Homo sapiens
 NAME/KEY: VARIANT
 LOCATION: (1) (2108)
 OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-854-856-30

Query Match 98.5%; Score 10646; DB 4; Length 2108;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 0; Indels 28; Gaps 1;
 1 MSGGAAGKOSTPGLFLSPAPAPKNGSSSDSVGKLGAAADAVTGRTEYRRRRHT 60
 1 MSGGAAGKOSTPGLFLSPAPAPKNGSSSDSVGKLGAAADAVTGRTEYRRRRHT 60
 61 MDKDSRGAATTTTTHRRFRFRSVICDSNAITALELPGLPLSLPQPSIPAAVQSAPEPH 120
 61 MDKDSRGAATTTTTHRRFRFRSVICDSNAITALELPGLPLSLPQPSIPAAVQSAPEPH 120
 121 REETVITATSOVAQOPPAAPAEQAVAGABSTVSSSKDPVQPSLVGSKKEPP 180
 121 REETVITATSOVAQOPPAAPAEQAVAGABSTVSSSKDPVQPSLVGSKKEPP 180
 181 ARSSGGGSAKEPQEBERSQOQDDELETFKAVGMSNGRFLKPIELIGRSGFKTVYGLD 240
 181 ARSSGGGSAKEPQEBERSQOQDDELETFKAVGMSNGRFLKPIELIGRSGFKTVYGLD 240
 241 TETVEAVACELQDRKLTSEORFKEBAEMKGLQHPNIVRFYDSWSTYKCKIVLV 300
 241 TETVEAVACELQDRKLTSEORFKEBAEMKGLQHPNIVRFYDSWSTYKCKIVLV 300
 301 TELMTSGTLTKYLRFRFMVKIKVLRSCROILKGLQFLHTTPTPIHRDLKCDNIFTGP 360
 301 TELMTSGTLTKYLRFRFMVKIKVLRSCROILKGLQFLHTTPTPIHRDLKCDNIFTGP 360
 361 TGSYKIGDLGATLKRAAFKSVIGTPEFMAPEMYEKEYESVVAFFGMCLMATASEY 420
 361 TGSYKIGDLGATLKRAAFKSVIGTPEFMAPEMYEKEYESVVAFFGMCLMATASEY 420
 421 PYSECQNAAQIYRRVTSGVKPAFDKVAIPVKEIIIEGCIKRONDERYSIDLNLHAFQ 480
 421 PYSECQNAAQIYRRVTSGVKPAFDKVAIPVKEIIIEGCIKRONDERYSIDLNLHAFQ 480
 481 BETGVARELAEDDEGEKIALMLIRIDIKLKQKYKONNEAIEFSPLDERVPEVDQEM 540
 481 BETGVARELAEDDEGEKIALMLIRIDIKLKQKYKONNEAIEFSPLDERVPEVDQEM 540
 541 VESGVCEGDHKTAKAIKDRVSLIKRKEORQVREOEKKQOESSLKQOVQSSASQ 600
 541 VESGVCEGDHKTAKAIKDRVSLIKRKEORQVREOEKKQOESSLKQOVQSSASQ 600
 601 TGIKOLPSASTGIPASTTSASVSTQVEPEPEADQHQLOQOQPSISVLSGTVDSQG 660
 601 TGIKOLPSASTGIPASTTSASVSTQVEPEPEADQHQLOQOQPSISVLSGTVDSQG 660
 661 SSVFSESVSSQOQTVSYSGOHEOASHSTGVGHIPTSTQAOQOPHGVYPPSSVAGQSGQ 720
 661 SSVFSESVSSQOQTVSYSGOHEOASHSTGVGHIPTSTQAOQOPHGVYPPSSVAGQSGQ 720

```

QY 721 QPSSSILTVSSSQPIQHPPQQQGIQOTAPPOQTQVSLSQSTSSSEATTAQPVSPQAP 780
Db 714 -----QQGIQOTAPPOQTQVSLSQSTSSSEATTAQPVSPQAP 752
QY 781 QVLPVNSAGKOSTQGVSVAPAEVAVAOPOATQPTTLASVDASHSVDASMSDGENV 840
Db 753 QVLPVNSAGKOSTQGVSVAPAEVAVAOPOATQPTTLASVDASHSVDASMSDGENV 812
QY 841 PSSGRHGRITTKHRYKRSVSRSHKETSRRPKILINVSNGKRVVECOLETHNRKAVT 900
Db 813 PSSGRHGRITTKHRYKRSVSRSHKETSRRPKILINVSNGKRVVECOLETHNRKAVT 872
QY 901 FKFDLDGNPEEIIATIMVNDPILAIERESFVDQVREIIEKADENLSDVSEPEPDQGL 960
Db 873 FKFDLDGNPEEIIATIMVNDPILAIERESFVDQVREIIEKADENLSDVSEPEPDQGL 932
QY 961 ESLQKDDYGGSGOKLEGEFKQPIPASMPQOIGIPSSLTQVVSAGRRFIVSPVES 1020
Db 933 ESLQKDDYGGSGOKLEGEFKQPIPASMPQOIGIPSSLTQVVSAGRRFIVSPVES 992
QY 1021 RLRESKVPSEITPTVAASTQSPGMNLSHASSLSLQOARSELPRADMEGPTAPENF 1080
Db 993 RLRESKVPSEITPTVAASTQSPGMNLSHASSLSLQOARSELPRADMEGPTAPENF 1052
QY 1081 SHTGPTPEPVVPPPLSIAAGVPTTAATAPVATSSPPMDISTSVYQSEVTVTEEGIAGV 1140
Db 1053 SHTGPTPEPVVPPPLSIAAGVPTTAATAPVATSSPPMDISTSVYQSEVTVTEEGIAGV 1112
QY 1141 ATSTGVVTSGLPLPEPVSESVLSSVSVSITIPAVVISITSSPELQVPTSEIIVSSTA 1200
Db 1113 ATSTGVVTSGLPLPEPVSESVLSSVSVSITIPAVVISITSSPELQVPTSEIIVSSTA 1172
QY 1201 LYPEVVSATSSASAGSTATPGPKPVPVVSQOAAAGSTVGTATLSVSTTSPFSTASQLS 1260
Db 1173 LYPEVVSATSSASAGSTATPGPKPVPVVSQOAAAGSTVGTATLSVSTTSPFSTASQLS 1232
QY 1261 IQLSSSTSTPLAATVVVSAHSLSDKTSHSSTTGLAFSLASAPSSSSGAGVSYISPGG 1320
Db 1233 IQLSSSTSTPLAATVVVSAHSLSDKTSHSSTTGLAFSLASAPSSSSGAGVSYISPGG 1292
QY 1321 LHPVTVSVASLSTPLILQOAGFTSTPLIPQVPSIPPLVQVAVANPAVQOQLIHNSQPAL 1380
Db 1293 LHPVTVSVASLSTPLILQOAGFTSTPLIPQVPSIPPLVQVAVANPAVQOQLIHNSQPAL 1352
QY 1381 LPNOPHTHCEVSDTQPKAPGIDITKLEKLSLSLSESSSSGAGASVLSFTSVIES 1440
Db 1353 LPNOPHTHCEVSDTQPKAPGIDITKLEKLSLSLSESSSSGAGASVLSFTSVIES 1412
QY 1441 TVTPGIPTTAAVAPSKLSTSTSTCLPPTNLPLGTVALPVTVPVPGQVSTPSTTSGVK 1500
Db 1413 TVTPGIPTTAAVAPSKLSTSTSTCLPPTNLPLGTVALPVTVPVPGQVSTPSTTSGVK 1472
QY 1501 PGTAPSKPPLTKAVLVEVTELPAGTLPSEQLPFPSPSLTQSOQPLEDDAQLRRTLSP 1560
Db 1473 PGTAPSKPPLTKAVLVEVTELPAGTLPSEQLPFPSPSLTQSOQPLEDDAQLRRTLSP 1532
QY 1561 EXITVTSVAVGVSMAAFTALTEAGTOPQKGVSVKKEBPVLTATSGAVFPMGRFQVSA 1620
Db 1533 EXITVTSVAVGVSMAAFTALTEAGTOPQKGVSVKKEBPVLTATSGAVFPMGRFQVSA 1592
QY 1621 DGAQKEGKNKSEDAKSVHFESSSTESSVLSSSSPESTLVKPEPNGITIPGISDVPESAH 1680
Db 1593 DGAQKEGKNKSEDAKSVHFESSSTESSVLSSSSPESTLVKPEPNGITIPGISDVPESAH 1652
QY 1681 KTTASEKASTQGTQPKVGRFOVTTANKVGRFVSYSKTEBDKITDTTKEGPAVAPPMDLEQ 1740
Db 1653 KTTASEKASTQGTQPKVGRFOVTTANKVGRFVSYSKTEBDKITDTTKEGPAVAPPMDLEQ 1712
QY 1741 AVLPVAVIPKKEKELESPSHLNGSSDPEAAFLSRVDVDSGSPHSHOJSSSLSPSONL 1800
Db 1713 AVLPVAVIPKKEKELESPSHLNGSSDPEAAFLSRVDVDSGSPHSHOJSSSLSPSONL 1772

```

```

QY 1801 SOSLSNFSNFSYMSDNESEDIEDIEDLKLELRIRDKHLKEIODLOSROKHIEISLYTKLG 1860
Db 1773 SOSLSNFSNFSYMSDNESEDIEDIEDLKLELRIRDKHLKEIODLOSROKHIEISLYTKLG 1832
QY 1861 KVPVAVIITPPAAPLSGRRRRPTKSGKSSRSSSLGKNSPOLSGNLSSQSAASYVHPQOT 1920
Db 1833 KVPVAVIITPPAAPLSGRRRRPTKSGKSSRSSSLGKNSPOLSGNLSSQSAASYVHPQOT 1892
QY 1921 LHPGNIPESGONLOPLKPSPSSDNLVSAFTSDGALISVPSLSAPGGGTSSTNTVGAIV 1980
Db 1893 LHPGNIPESGONLOPLKPSPSSDNLVSAFTSDGALISVPSLSAPGGGTSSTNTVGAIV 1952
QY 1981 NSQAAQOPPAMTSRRKGTFTDHLKLVDMNARMNLSGRRSGKGNVYEGPGARKFS 2040
Db 1953 NSQAAQOPPAMTSRRKGTFTDHLKLVDMNARMNLSGRRSGKGNVYEGPGARKFS 2012
QY 2041 APQOLCISMTSNLGSAPISASATSLGHTFKSMCPQOQYFPATPFQAGMSGTGAPAPQ 2100
Db 2013 APQOLCISMTSNLGSAPISASATSLGHTFKSMCPQOQYFPATPFQAGMSGTGAPAPQ 2072
QY 2101 PLGQFQVGTASLQNFNINISNLOKSIINPPGSLNRTT 2136
Db 2073 PLGQFQVGTASLQNFNINISNLOKSIINPPGSLNRTT 2108

RESULT 5
US-09-854-856-24
Sequence 24, Application US/09854856
Patent No. 6541252
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Halbur, Erin
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
FILE REFERENCE: LEX-0178-usa
CURRENT APPLICATION NUMBER: US/09/854,856
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: US 60/206,015
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 2201
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)-(2201)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-24

Query Match 97.9%; Score 10589.5; DB 4; Length 2201;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 2108; Conservative 0; Mismatches 0; Indels 121; Gaps 2;

```

QY 241 TETTVAVACELQORUKTKSEBROFKEBAEMKGLQHPNIVRFYDSESTVKGKCIIVL 300
 Db 241 TETTVAVACELQORUKTKSEBROFKEBAEMKGLQHPNIVRFYDSESTVKGKCIIVL 300
 QY 301 TELMTSTLTKYLRFRFVWKIKVLRSCROILKGLQHLHRTPIIHRDLKCDNIPTTGP 360
 Db 301 TELMTSTLTKYLRFRFVWKIKVLRSCROILKGLQHLHRTPIIHRDLKCDNIPTTGP 360
 QY 361 TGSVKIGDLGLATKRAFSAKSVIGTEPEFMAPEWYEKEKDESVVVAFGMCMLMATESEY 420
 Db 361 TGSVKIGDLGLATKRAFSAKSVIGTEPEFMAPEWYEKEKDESVVVAFGMCMLMATESEY 420
 QY 421 PYSECQNAAQIYRRVTSVGVKRPASFDKVAIPEVKEIIEGCIJRONKDEBYSIKDLNHAFFQ 480
 Db 421 PYSECQNAAQIYRRVTSVGVKRPASFDKVAIPEVKEIIEGCIJRONKDEBYSIKDLNHAFFQ 480
 QY 481 EETGVARELAEBEDGEXIAIKLWLRIDIKLKKKYKDNDAIESFDERVPEVDVAQEM 540
 Db 481 EETGVARELAEBEDGEXIAIKLWLRIDIKLKKKYKDNDAIESFDERVPEVDVAQEM 540
 QY 541 VESGVCEGDKHTAKAKIKDRVSLIKRRREOROLVREBQEKKESSSLKQOVQSSASQ 600
 Db 541 VESGVCEGDKHTAKAKIKDRVSLIKRRREOROLVREBQEKKESSSLKQOVQSSASQ 600
 QY 601 TGIKQLPASATGIPFASSTSAVSSTQVPEBPEADQHQLOQYQOPISVLSGTVDSGQG 660
 Db 601 TGIKQLPASATGIPFASSTSAVSSTQVPEBPEADQHQLOQYQOPISVLSGTVDSGQG 660
 QY 661 SSVTERSVSSQOVVSGSHQHOAHSTGYVPGHIPSIVQAOOSOPHGYPPSSVAQOSQSG 720
 Db 661 SSVTERSVSSQOVVSGSHQHOAHSTGYVPGHIPSIVQAOOSOPHGYPPSSVAQOSQSG 720
 QY 721 QPSSSITLVSSSQPIQHPQQQOGIIOQTAPPOQTVOVQSLSTQTSSEATTQOPPOAP 780
 Db 721 QPSSSITLVSSSQPIQHPQQQOGIIOQTAPPOQTVOVQSLSTQTSSEATTQOPPOAP 780
 QY 781 QVLPVSAAGKQ-----QOGIIOQTAPPOQTVOVQSLSTQTSSEATTQOPPOAP 791
 Db 781 QVLPVSAAGKQ-----QOGIIOQTAPPOQTVOVQSLSTQTSSEATTQOPPOAP 791
 QY 792 QVLPVSAAGKQ-----QOGIIOQTAPPOQTVOVQSLSTQTSSEATTQOPPOAP 812
 Db 792 QVLPVSAAGKQ-----QOGIIOQTAPPOQTVOVQSLSTQTSSEATTQOPPOAP 812
 QY 813 VVOQPVESNLLVPNGVGGQYVVSQPGSGLAQAPTTSSQOAVLBSSTQGVSAVAEPAVAV 872
 Db 813 VVOQPVESNLLVPNGVGGQYVVSQPGSGLAQAPTTSSQOAVLBSSTQGVSAVAEPAVAV 872
 QY 808 AOPATOPPTLASSVDSADSVASGMSDGNENVPSSGSRHGRITTKXHYRKSVSRSRHE 867
 Db 808 AOPATOPPTLASSVDSADSVASGMSDGNENVPSSGSRHGRITTKXHYRKSVSRSRHE 867
 QY 873 AOPATOPPTLASSVDSADSVASGMSDGNENVPSSGSRHGRITTKXHYRKSVSRSRHE 932
 Db 873 AOPATOPPTLASSVDSADSVASGMSDGNENVPSSGSRHGRITTKXHYRKSVSRSRHE 932
 QY 868 KTSRPRKRLINLVSNKGRVVECOLETHNRKAVTEFKFLDGDNPBEIATIMVNDPFLAIE 927
 Db 868 KTSRPRKRLINLVSNKGRVVECOLETHNRKAVTEFKFLDGDNPBEIATIMVNDPFLAIE 927
 QY 933 KTSRPRKRLINLVSNKGRVVECOLETHNRKAVTEFKFLDGDNPBEIATIMVNDPFLAIE 992
 Db 933 KTSRPRKRLINLVSNKGRVVECOLETHNRKAVTEFKFLDGDNPBEIATIMVNDPFLAIE 992
 QY 928 RESFYDOVREIIEKADDEMLSEDVVEPEBGDGLSLQKDDYVSGSGOKLEGEFKOPIPA 987
 Db 928 RESFYDOVREIIEKADDEMLSEDVVEPEBGDGLSLQKDDYVSGSGOKLEGEFKOPIPA 987
 QY 993 RESFYDOVREIIEKADDEMLSEDVVEPEBGDGLSLQKDDYVSGSGOKLEGEFKOPIPA 1052
 Db 993 RESFYDOVREIIEKADDEMLSEDVVEPEBGDGLSLQKDDYVSGSGOKLEGEFKOPIPA 1052
 QY 988 SSMPOQOIGIPPTSSITLOVYHSAGRFPIVSPVESRLRESKVPSEITPDIVASTQSGMN 1047
 Db 988 SSMPOQOIGIPPTSSITLOVYHSAGRFPIVSPVESRLRESKVPSEITPDIVASTQSGMN 1047
 QY 1053 SSMPOQOIGIPPTSSITLOVYHSAGRFPIVSPVESRLRESKVPSEITPDIVASTQSGMN 1112
 Db 1053 SSMPOQOIGIPPTSSITLOVYHSAGRFPIVSPVESRLRESKVPSEITPDIVASTQSGMN 1112
 QY 1048 LSHSASSLSLQOAFSELRRAQMTGEPNTAPNFSHTGPTFPVVPFLSSIAAGVPTTAAT 1107
 Db 1048 LSHSASSLSLQOAFSELRRAQMTGEPNTAPNFSHTGPTFPVVPFLSSIAAGVPTTAAT 1107
 QY 1113 LSHSASSLSLQOAFSELRRAQMTGEPNTAPNFSHTGPTFPVVPFLSSIAAGVPTTAAT 1172
 Db 1113 LSHSASSLSLQOAFSELRRAQMTGEPNTAPNFSHTGPTFPVVPFLSSIAAGVPTTAAT 1172
 QY 1108 APVATATSPNDISTSVIQSEVTVPTBEGIANATSTCVVTSGLPIPIPVSESPLYSSV 1167
 Db 1108 APVATATSPNDISTSVIQSEVTVPTBEGIANATSTCVVTSGLPIPIPVSESPLYSSV 1167
 QY 1173 APVATATSPNDISTSVIQSEVTVPTBEGIANATSTCVVTSGLPIPIPVSESPLYSSV 1232
 Db 1173 APVATATSPNDISTSVIQSEVTVPTBEGIANATSTCVVTSGLPIPIPVSESPLYSSV 1232
 QY 1168 SSITIPAVSISTSPSLQVPTSTSEIIVASTALYPSVTAVTSASAGSSTATGPPXPA 1227
 Db 1168 SSITIPAVSISTSPSLQVPTSTSEIIVASTALYPSVTAVTSASAGSSTATGPPXPA 1227
 QY 1233 SSITIPAVSISTSPSLQVPTSTSEIIVASTALYPSVTAVTSASAGSSTATGPPXPA 1232
 Db 1233 SSITIPAVSISTSPSLQVPTSTSEIIVASTALYPSVTAVTSASAGSSTATGPPXPA 1232
 QY 1228 VVSQOAAAGSTTVGATLTVSVSTTSPSTASQSLISQSSSTSTPFLAETVVVSAHSLDKTS 1287
 Db 1228 VVSQOAAAGSTTVGATLTVSVSTTSPSTASQSLISQSSSTSTPFLAETVVVSAHSLDKTS 1287

Db 1293 VVSQOAAAGSTTVGATLTVSVSTTSPSTASQSLISQSSSTSTPFLAETVVVSAHSLDKTS 1352
 QY 1288 HSSITGLAFSLAPSSSSSPGAVSSYISOPGGIHPVIVSVASTPILPOAAPTSGNPL 1347
 Db 1288 HSSITGLAFSLAPSSSSSPGAVSSYISOPGGIHPVIVSVASTPILPOAAPTSGNPL 1347
 QY 1353 HSSITGLAFSLAPSSSSSPGAVSSYISOPGGIHPVIVSVASTPILPOAAPTSGNPL 1412
 Db 1353 HSSITGLAFSLAPSSSSSPGAVSSYISOPGGIHPVIVSVASTPILPOAAPTSGNPL 1412
 QY 1348 LPQVPSIPLVQPVANVPAVQOTLIHSQOPALLPNOPTHCEVDSOTPKAAGIDIK 1407
 Db 1348 LPQVPSIPLVQPVANVPAVQOTLIHSQOPALLPNOPTHCEVDSOTPKAAGIDIK 1407
 QY 1413 LPQVPSIPLVQPVANVPAVQOTLIHSQOPALLPNOPTHCEVDSOTPKAAGIDIK 1472
 Db 1413 LPQVPSIPLVQPVANVPAVQOTLIHSQOPALLPNOPTHCEVDSOTPKAAGIDIK 1472
 QY 1408 TLEBKRLSLESHSSGCAVASVLETSVLESTVTPGIPITTVAPASKLIITSTTCLPP 1467
 Db 1408 TLEBKRLSLESHSSGCAVASVLETSVLESTVTPGIPITTVAPASKLIITSTTCLPP 1467
 QY 1473 TLEBKRLSLESHSSGCAVASVLETSVLESTVTPGIPITTVAPASKLIITSTTCLPP 1532
 Db 1473 TLEBKRLSLESHSSGCAVASVLETSVLESTVTPGIPITTVAPASKLIITSTTCLPP 1532
 QY 1468 TNLPIGTVALPVTPVTPGQVSTPSTSTTSGVGRPAKSPKPLTKAVLVGTELPAGTL 1527
 Db 1468 TNLPIGTVALPVTPVTPGQVSTPSTSTTSGVGRPAKSPKPLTKAVLVGTELPAGTL 1527
 QY 1533 TNLPIGTVALPVTPVTPGQVSTPSTSTTSGVGRPAKSPKPLTKAVLVGTELPAGTL 1592
 Db 1533 TNLPIGTVALPVTPVTPGQVSTPSTSTTSGVGRPAKSPKPLTKAVLVGTELPAGTL 1592
 QY 1528 PSEQLPFPGPSLTSQOPLLEDLDAQLRLTSEKXITVTSAVGVSAAPATAITEAGTOP 1587
 Db 1528 PSEQLPFPGPSLTSQOPLLEDLDAQLRLTSEKXITVTSAVGVSAAPATAITEAGTOP 1587
 QY 1593 PSEQLPFPGPSLTSQOPLLEDLDAQLRLTSEKXITVTSAVGVSAAPATAITEAGTOP 1652
 Db 1593 PSEQLPFPGPSLTSQOPLLEDLDAQLRLTSEKXITVTSAVGVSAAPATAITEAGTOP 1652
 QY 1588 QKGVSOYKEGPVATSSGAGVFYKGRFOVVAADGAQKBNKSEDAKSVHESSTSSS 1647
 Db 1588 QKGVSOYKEGPVATSSGAGVFYKGRFOVVAADGAQKBNKSEDAKSVHESSTSSS 1647
 QY 1653 QKGVSOYKEGPVATSSGAGVFYKGRFOVVAADGAQKBNKSEDAKSVHESSTSSS 1712
 Db 1653 QKGVSOYKEGPVATSSGAGVFYKGRFOVVAADGAQKBNKSEDAKSVHESSTSSS 1712
 QY 1648 VLSSSSPSTLTVPEPFGITIPGISDVPEPSAKHTTSEAKSDPTGQTKVGRFOVTTAN 1707
 Db 1648 VLSSSSPSTLTVPEPFGITIPGISDVPEPSAKHTTSEAKSDPTGQTKVGRFOVTTAN 1707
 QY 1713 VLSSSSPSTLTVPEPFGITIPGISDVPEPSAKHTTSEAKSDPTGQTKVGRFOVTTAN 1772
 Db 1713 VLSSSSPSTLTVPEPFGITIPGISDVPEPSAKHTTSEAKSDPTGQTKVGRFOVTTAN 1772
 QY 1708 KVGRFSYKTEDKLTIDTKKEGPVAPPEFMDLEQAVLPAVLPKKEPELSEPSHINCESD 1767
 Db 1708 KVGRFSYKTEDKLTIDTKKEGPVAPPEFMDLEQAVLPAVLPKKEPELSEPSHINCESD 1767
 QY 1773 KVGRFSYKTEDKLTIDTKKEGPVAPPEFMDLEQAVLPAVLPKKEPELSEPSHINCESD 1832
 Db 1773 KVGRFSYKTEDKLTIDTKKEGPVAPPEFMDLEQAVLPAVLPKKEPELSEPSHINCESD 1832
 QY 1768 PEAAFLSRDVEDDGGSGPHSHQULSKSLPEQNTSOSISNFSNYSNDSNEDIEDDLK 1827
 Db 1768 PEAAFLSRDVEDDGGSGPHSHQULSKSLPEQNTSOSISNFSNYSNDSNEDIEDDLK 1827
 QY 1833 PEAAFLSRDVEDDGGSGPHSHQULSKSLPEQNTSOSISNFSNYSNDSNEDIEDDLK 1892
 Db 1833 PEAAFLSRDVEDDGGSGPHSHQULSKSLPEQNTSOSISNFSNYSNDSNEDIEDDLK 1892
 QY 1828 LEIARLRDKHLEIODLOSQKHIEISLYRKLGVPAVAVIIPPAFLSGRRRRRTKSGS 1887
 Db 1828 LEIARLRDKHLEIODLOSQKHIEISLYRKLGVPAVAVIIPPAFLSGRRRRRTKSGS 1887
 QY 1893 LEIARLRDKHLEIODLOSQKHIEISLYRKLGVPAVAVIIPPAFLSGRRRRRTKSGS 1952
 Db 1893 LEIARLRDKHLEIODLOSQKHIEISLYRKLGVPAVAVIIPPAFLSGRRRRRTKSGS 1952
 QY 1888 KSSRSSLSLGNKSPDLSGNLSGOSNAAYLHPQOTLHPGNIPESGONOLOPLKPSPSDN 1947
 Db 1888 KSSRSSLSLGNKSPDLSGNLSGOSNAAYLHPQOTLHPGNIPESGONOLOPLKPSPSDN 1947
 QY 1953 KSSRSSLSLGNKSPDLSGNLSGOSNAAYLHPQOTLHPGNIPESGONOLOPLKPSPSDN 2012
 Db 1953 KSSRSSLSLGNKSPDLSGNLSGOSNAAYLHPQOTLHPGNIPESGONOLOPLKPSPSDN 2012
 QY 1948 LYSAFITSDGAIYVPSLSAPQGTISNTTVGATVNSQAQAQOPAMTSSRKGTFTDDJLKL 2007
 Db 1948 LYSAFITSDGAIYVPSLSAPQGTISNTTVGATVNSQAQAQOPAMTSSRKGTFTDDJLKL 2007
 QY 2013 LYSAFITSDGAIYVPSLSAPQGTISNTTVGATVNSQAQAQOPAMTSSRKGTFTDDJLKL 2072
 Db 2013 LYSAFITSDGAIYVPSLSAPQGTISNTTVGATVNSQAQAQOPAMTSSRKGTFTDDJLKL 2072
 QY 2008 VDNWAPRANMLSGRSGKGMNTEGPGMAKFFSAPQGLCTSMNLSGSAPIASAATSL 2067
 Db 2008 VDNWAPRANMLSGRSGKGMNTEGPGMAKFFSAPQGLCTSMNLSGSAPIASAATSL 2067
 QY 2073 VDNWAPRANMLSGRSGKGMNTEGPGMAKFFSAPQGLCTSMNLSGSAPIASAATSL 2132
 Db 2073 VDNWAPRANMLSGRSGKGMNTEGPGMAKFFSAPQGLCTSMNLSGSAPIASAATSL 2132
 QY 2068 GHFTKSMCPQOYGFAPTPFGAOWSGTGAPAPLOLGFQFVGTASLQNFNINSLOKSISN 2127
 Db 2068 GHFTKSMCPQOYGFAPTPFGAOWSGTGAPAPLOLGFQFVGTASLQNFNINSLOKSISN 2127
 QY 2133 GHFTKSMCPQOYGFAPTPFGAOWSGTGAPAPLOLGFQFVGTASLQNFNINSLOKSISN 2192
 Db 2133 GHFTKSMCPQOYGFAPTPFGAOWSGTGAPAPLOLGFQFVGTASLQNFNINSLOKSISN 2192
 QY 2128 PPGSNLRTT 2136
 Db 2128 PPGSNLRTT 2136
 QY 2193 PPGSNLRTT 2201
 Db 2193 PPGSNLRTT 2201

RESULT 6
 US-09-854-856-18
 / Sequence 18, Application US/09854856
 / Patent No. 6541252
 / GENERAL INFORMATION:
 / APPLICANT: Walke, D. Wade
 / APPLICANT: Hilbur, Erin
 / APPLICANT: Donoho, Gregory
 / APPLICANT: Turner, C. Alexander Jr.

TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
FILE REFERENCE: Encoding the Same
CURRENT APPLICATION NUMBER: US/09/854,856
CURRENT FILING DATE: 2001-05-14
PRIORITY APPLICATION NUMBER: US 60/206,015
PRIORITY FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 18
LENGTH: 2354
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(2354)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-18

Query Match 97.2%; Score 10513; DB 4; Length 2354;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 2108; Conservative 0; Mismatches 0; Indels 274; Gaps 2;

QY 1 MSGAAEKOSTPGSLFLSPAPAPKNGSSDSSVGEKLGAAAADAVTGRTEYRRRRHT 60
DB 1 MSGAAEKOSTPGSLFLSPAPAPKNGSSDSSVGEKLGAAAADAVTGRTEYRRRRHT 60
QY 61 MDKSRGAAATTTTTHRRFFRRSVICDNNATALLPLPLPSPSTPAAPVQSAPEPH 120
DB 61 MDKSRGAAATTTTTHRRFFRRSVICDNNATALLPLPLPSPSTPAAPVQSAPEPH 120
QY 121 REETVATATSOVAQPPAAAAPGEQAVAPAPSTVSTSKDRPVQPSLVGSKPEPP 180
DB 121 REETVATATSOVAQPPAAAAPGEQAVAPAPSTVSTSKDRPVQPSLVGSKPEPP 180
QY 121 REETVATATSOVAQPPAAAAPGEQAVAPAPSTVSTSKDRPVQPSLVGSKPEPP 180
DB 121 REETVATATSOVAQPPAAAAPGEQAVAPAPSTVSTSKDRPVQPSLVGSKPEPP 180
QY 181 ARSGGGGSAKEPOEERSQOQDDIEELRYAVGMSNDGRFLKPDIEIGRGSFKTVKGLD 240
DB 181 ARSGGGGSAKEPOEERSQOQDDIEELRYAVGMSNDGRFLKPDIEIGRGSFKTVKGLD 240
QY 241 TETVEVAMCELODRKLTCSRORFKEEAEMKGLQHPNIVRFDSMESTYKGGKCVLV 300
DB 241 TETVEVAMCELODRKLTCSRORFKEEAEMKGLQHPNIVRFDSMESTYKGGKCVLV 300
QY 301 TELMTSGTLTYLKRFPYMKIKVLRSMCRQILKGLQFLHTPTPIIRHDLKCDNIFITGP 360
DB 301 TELMTSGTLTYLKRFPYMKIKVLRSMCRQILKGLQFLHTPTPIIRHDLKCDNIFITGP 360
QY 301 TELMTSGTLTYLKRFPYMKIKVLRSMCRQILKGLQFLHTPTPIIRHDLKCDNIFITGP 360
DB 301 TELMTSGTLTYLKRFPYMKIKVLRSMCRQILKGLQFLHTPTPIIRHDLKCDNIFITGP 360
QY 361 TGSVKIGDLGLATLKRAFAKSVIGTPEFMAPEMYEKEYDESVDVYAFGMCMEMATSEY 420
DB 361 TGSVKIGDLGLATLKRAFAKSVIGTPEFMAPEMYEKEYDESVDVYAFGMCMEMATSEY 420
QY 421 PYSGCQAAQIYRRVTSVGVKSPASFDKVAIPVEKEIIGCIRONDDEYYSIKDLINAFQ 480
DB 421 PYSGCQAAQIYRRVTSVGVKSPASFDKVAIPVEKEIIGCIRONDDEYYSIKDLINAFQ 480
QY 481 EETGVRELAEDDGEKIAIKMLRIDIKLKAKYKONENAIERSFLERDVEDVAQEM 540
DB 481 EETGVRELAEDDGEKIAIKMLRIDIKLKAKYKONENAIERSFLERDVEDVAQEM 540
QY 541 VESGYVEGHHKMAKAIKDRVSLIKRRRQRLVREOEKKKQESSLKQVEQSSASQ 600
DB 541 VESGYVEGHHKMAKAIKDRVSLIKRRRQRLVREOEKKKQESSLKQVEQSSASQ 600
QY 601 TGTGOLPSASTGPTASTASVSTOVEPEPEADQHQQLQYQPSISVSDGTVDGSG 660
DB 601 TGTGOLPSASTGPTASTASVSTOVEPEPEADQHQQLQYQPSISVSDGTVDGSG 660
QY 661 SSYFTESRVSSQOQVTSYGSQHEQASHGTGVPGHIPSTVQAQSOQPHGYVPSSVAAQSGG 720
DB 661 SSYFTESRVSSQOQVTSYGSQHEQASHGTGVPGHIPSTVQAQSOQPHGYVPSSVAAQSGG 720
QY 721 QPSSSSITGVSSQPIQHPQOQGIQCTAPPOQTVYSLQSTSSSEATTAQPVQPOAP 780
DB 721 QPSSSSITGVSSQPIQHPQOQGIQCTAPPOQTVYSLQSTSSSEATTAQPVQPOAP 780

DB 714 -----QCGIQQCTAPPOQTVYSLQSTSSSEATTAQPVQPOAP 752
QY 781 QVLPVSAKQ----- 791
DB 753 QVLPVSAKQLPVQCPVPTIGBPQIPVATQPSVVPVSHAHFLPVQPLPTLLPOYP 812
QY 792 ----- 791
DB 813 VSQIPITPHVSTACQGBSSLPITMAAGITQPLLTASSATTAAIPGVSTVPSQLPTLL 872
QY 792 ----- 791
DB 873 QPVTQLPSPVHPQLQPAVQSMGIPANLQGAEEVPLSSGDVLXQGFPRRLPQYPGDSNI 932
QY 792 ----- 791
DB 933 APSNNAVACIHSTVLXPMPEVTLATPGYPTVQPYVESNLLVPMGVGVQVQVQSG 992
QY 792 -----STQGVQVAPAEVAVVQAPQATQPTTLASSVDSAHSDVASGMS 834
DB 993 GSIAQATTSQCAVLESTQGVQVAPAEVAVVQAPQATQPTTLASSVDSAHSDVASGMS 1052
QY 835 DGNENVSSGRHGRITTKHRYKSVSRSHKXTSRPKLILNVSNKGRVVECOLETH 894
DB 1053 DGNENVSSGRHGRITTKHRYKSVSRSHKXTSRPKLILNVSNKGRVVECOLETH 1112
QY 895 NRKVTTFKPLDGNPEEITATIMNNDPILAEBSVVDYREILEKADMLSDVSEY 954
DB 1113 NRKVTTFKPLDGNPEEITATIMNNDPILAEBSVVDYREILEKADMLSDVSEY 1172
QY 955 EGDQGLSLQKNDYGVSGSKLEGEFKQPIPASMSQOIGIPSSSLQVYVSHSAGRPTV 1014
DB 1173 EGDQGLSLQKNDYGVSGSKLEGEFKQPIPASMSQOIGIPSSSLQVYVSHSAGRPTV 1222
QY 1015 SPVSESLRESKVPSEITDTVAASTAQSPPQMLSHSASSLSLQAFSELRRAQMTGPN 1074
DB 1233 SPVSESLRESKVPSEITDTVAASTAQSPPQMLSHSASSLSLQAFSELRRAQMTGPN 1292
QY 1075 TAPNPSHTGPTTPVPPPELSSIAQVPTTAAATPVPATSSPPNDISTSVYQSVTYPT 1134
DB 1293 TAPNPSHTGPTTPVPPPELSSIAQVPTTAAATPVPATSSPPNDISTSVYQSVTYPT 1352
QY 1135 EGIAGVASTGVYSGGLPIPPSESEPVLSVSSSITIPVAVSISTSPSLOVTSSEI 1194
DB 1353 EGIAGVASTGVYSGGLPIPPSESEPVLSVSSSITIPVAVSISTSPSLOVTSSEI 1412
QY 1195 VVSSSTALYPSVTVSATSASAGSSTATPGPKPAPVAVSQAAAGSTTVGATLTLSVSTTSPS 1254
DB 1413 VVSSSTALYPSVTVSATSASAGSSTATPGPKPAPVAVSQAAAGSTTVGATLTLSVSTTSPS 1472
QY 1255 TAOQLSTQLSSSTPTLATTVVVAHSLDKTSHSSTTGIAFLSAPSSSSPQAGVSSY 1314
DB 1473 TAOQLSTQLSSSTPTLATTVVVAHSLDKTSHSSTTGIAFLSAPSSSSPQAGVSSY 1532
QY 1315 ISOQGLHPVLPVSVIASTPILPQAAGPTSTPLLPQVPSIPVQVPAANPAVQOQLIHS 1374
DB 1533 ISOQGLHPVLPVSVIASTPILPQAAGPTSTPLLPQVPSIPVQVPAANPAVQOQLIHS 1552
QY 1375 QPOPALIPNOPTHCPVSDSTQPKAPGIDDKTEBKLSLFSHSSGAQHASVLET 1434
DB 1553 QPOPALIPNOPTHCPVSDSTQPKAPGIDDKTEBKLSLFSHSSGAQHASVLET 1652
QY 1435 SLVIESVTPGIPPTVAVPSKLTSTSTCLPPTNPLPGVVALPVMPVYVPGQVSTPVST 1494
DB 1653 SLVIESVTPGIPPTVAVPSKLTSTSTCLPPTNPLPGVVALPVMPVYVPGQVSTPVST 1712
QY 1495 TTSQVKGTPASKRPPLTKAVLPLVGTLPSEQLPSEQLPFPQPSLQSQOQPLEDDAQL 1554
DB 1713 TTSQVKGTPASKRPPLTKAVLPLVGTLPSEQLPSEQLPFPQPSLQSQOQPLEDDAQL 1772
QY 1555 RRTLSPEXITTVTSAGVSVMAAPTAITTEAGTQPOKGVSYQVKEGVALTSSGAGVFXKGRF 1614
DB 1773 RRTLSPEXITTVTSAGVSVMAAPTAITTEAGTQPOKGVSYQVKEGVALTSSGAGVFXKGRF 1832

QY 1615 QVSVAADGAQKGNKSEDAKSVHFEESSTSESSVLSSSSPESSTLYKEEPNGITIPGISSD 1674
DB 1833 QVSVAADGAQKGNKSEDAKSVHFEESSTSESSVLSSSSPESSTLYKEEPNGITIPGISSD 1892
QY 1675 VPESAHHTTSAEASDITGQPTKVRFOYTTANKVGRFSVSKTEDKLTDTKKEGPVAPSP 1734
DB 1693 VPESAHHTTSAEASDITGQPTKVRFOYTTANKVGRFSVSKTEDKLTDTKKEGPVAPSP 1952
QY 1735 FMDLEQAVLPVAVIIPKKEKPELSEPSHNGPESDPEAFLSRDNDVDSGSPSPHQLSKS 1794
DB 1953 FMDLEQAVLPVAVIIPKKEKPELSEPSHNGPESDPEAFLSRDNDVDSGSPSPHQLSKS 2012
QY 1795 LPSQNTLSQSLNSFNSSNYSMSDNEDEDEDLKLELRRLDKHKEIODLQSRKHIES 1854
DB 2013 LPSQNTLSQSLNSFNSSNYSMSDNEDEDEDLKLELRRLDKHKEIODLQSRKHIES 2072
QY 1855 LYTGLGVPPAVIIPPAAPLUSGRRRRPTKSGKSSSSSISGKNSPOLSGNLSGQSAASV 1914
DB 2073 LYTGLGVPPAVIIPPAAPLUSGRRRRPTKSGKSSSSSISGKNSPOLSGNLSGQSAASV 2132
QY 1915 LHPQQLHPGNIPESSQNLQPLKPSPSDNLYSAFSTSGATSVPSLSAPGQGTSTN 1974
DB 2133 LHPQQLHPGNIPESSQNLQPLKPSPSDNLYSAFSTSGATSVPSLSAPGQGTSTN 2192
QY 1975 TVGATVNSQAAQAPPMATSSRKCTFTDHLKLVDMNARDAMNLSGRGSGKHNNYEGPG 2034
DB 2193 TVGATVNSQAAQAPPMATSSRKCTFTDHLKLVDMNARDAMNLSGRGSGKHNNYEGPG 2252
QY 2035 MARKFSAPGOLCISMTSNLGGASAPISAAASATSLGHFTKSCWCPQOYGPATPGQWNSGT 2094
DB 2253 MARKFSAPGOLCISMTSNLGGASAPISAAASATSLGHFTKSCWCPQOYGPATPGQWNSGT 2312
QY 2095 GGAPQPLGQFQPVGTASLQNFENISNLOKISNPPGSLKRTT 2136
DB 2313 GGAPQPLGQFQPVGTASLQNFENISNLOKISNPPGSLKRTT 2354

RESULT 7

US-09-854-856-46
Sequence 46, Application US/09854856
Patent No. 6541252
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Hilbun, Erin
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
FILE REFERENCE: Lex-0178-USA
CURRENT APPLICATION NUMBER: US/09/854, 856
CURRENT FILING DATE: 2001-05-14
PRIORITY APPLICATION NUMBER: US 60/206, 015
PRIORITY FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 46
LENGTH: 2076
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(2076)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-46

Query Match 97.2%, Score 10508, DB 4, Length 2076,
Best Local Similarity 100.0%, Pred. No. 0,
Matches 2076, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 61 MDKSRGAATTTTTHFRFRSVICDSNATALELPGPLSLPQSPSTPAAPQSAAPPEPH 120
DB 1 MDKSRGAATTTTTHFRFRSVICDSNATALELPGPLSLPQSPSTPAAPQSAAPPEPH 60

QY 121 REETVATATATSOVAQOQPAAABEGQAVAGPASTVPSSTSKDRPVSQPSLVGSKKEPP 180
DB 61 REETVATATATSOVAQOQPAAABEGQAVAGPASTVPSSTSKDRPVSQPSLVGSKKEPP 120
QY 181 ARSGSGGGAKEPQEBRSQOQDDIIELETAIVKMSNDGRFLKDEIFGRGSPKTVYGLD 240
DB 121 ARSGSGGGAKEPQEBRSQOQDDIIELETAIVKMSNDGRFLKDEIFGRGSPKTVYGLD 180
QY 241 TETTVAVAMELODRKLTSEORPFKEAEMLGLQHPNIVRFDSNESTVKGKCVLV 300
DB 181 TETTVAVAMELODRKLTSEORPFKEAEMLGLQHPNIVRFDSNESTVKGKCVLV 240
QY 301 TELMTSGTLTYLKRPFVWKIKYLRSWCRQILKGLQFLHTRTPPIIHRDICKDNIFITGP 360
DB 241 TELMTSGTLTYLKRPFVWKIKYLRSWCRQILKGLQFLHTRTPPIIHRDICKDNIFITGP 300
QY 361 TGSVKIGDLGLATLKRASPAKSVYGTPEFMAPEMEYEEKYDSVDVAFGCMLEMATSEY 420
DB 301 TGSVKIGDLGLATLKRASPAKSVYGTPEFMAPEMEYEEKYDSVDVAFGCMLEMATSEY 360
QY 421 PYSECQNAQIYRRVTSGVAPSPDKVAIPEVKEIIEGCIKONDERYSIKDLNHAFFQ 480
DB 361 PYSECQNAQIYRRVTSGVAPSPDKVAIPEVKEIIEGCIKONDERYSIKDLNHAFFQ 420
QY 481 BETGVRELAEBEDGEXIAIKLWLRJEDIKLKGXYKDNABIESFDLERDVEDVAQEM 540
DB 421 BETGVRELAEBEDGEXIAIKLWLRJEDIKLKGXYKDNABIESFDLERDVEDVAQEM 480
QY 541 VESGIVCEGDKHTMAKAIKORVSLIKRRERQOLVREDEKKKQOEBESLKQOYBOSSAQ 600
DB 481 VESGIVCEGDKHTMAKAIKORVSLIKRRERQOLVREDEKKKQOEBESLKQOYBOSSAQ 540
QY 601 TGIKOLPASTGIPATSTASVSTQVEPEPEADHQOLQYOQPSISVSDGVDSGQG 660
DB 541 TGIKOLPASTGIPATSTASVSTQVEPEPEADHQOLQYOQPSISVSDGVDSGQG 600
QY 661 SSVFTEBSRVSSQOQVSYSGQHEQAHSTGTYPGHIPTVQAQSQPHGVYPPSSVVAQSQSQ 720
DB 601 SSVFTEBSRVSSQOQVSYSGQHEQAHSTGTYPGHIPTVQAQSQPHGVYPPSSVVAQSQSQ 660
QY 721 QPSSSLTGVSSSQPIHPQOQOQIGQTPAPQOQVSYLSQSTSSSEATTAQVPSQDAP 780
DB 661 QPSSSLTGVSSSQPIHPQOQOQIGQTPAPQOQVSYLSQSTSSSEATTAQVPSQDAP 720
QY 781 QVLPQVAGKQSTQGVQVAPAEVAVAOQATQPTLLASVDSABSDVAGSGMDGENV 840
DB 721 QVLPQVAGKQSTQGVQVAPAEVAVAOQATQPTLLASVDSABSDVAGSGMDGENV 780
QY 841 PSSSGRHEGRITTKHRYKSVRSRHEKTSRPKLILVNSNKGDRVVECOLETHNRKVT 900
DB 781 PSSSGRHEGRITTKHRYKSVRSRHEKTSRPKLILVNSNKGDRVVECOLETHNRKVT 840
QY 901 FKPLDQNPBEIATIVNNDPILAIRESFVQVRIIEKADMLESDVSVREBQGL 960
DB 841 FKPLDQNPBEIATIVNNDPILAIRESFVQVRIIEKADMLESDVSVREBQGL 900
QY 961 ESLQKXDYDPSGSGQKEGEKQPIPASMPQOIGITSSLTQVYHAGRRFVSPVES 1020
DB 901 ESLQKXDYDPSGSGQKEGEKQPIPASMPQOIGITSSLTQVYHAGRRFVSPVES 960
QY 1021 RLRESKVPSEBITDTVAASTAQSFGNMLSHSASSLSQQAFFSELRRQOMTEGPTAPNF 1080
DB 961 RLRESKVPSEBITDTVAASTAQSFGNMLSHSASSLSQQAFFSELRRQOMTEGPTAPNF 1020
QY 1081 SHTGPTFPVVPVPPFSSITAGVPTTAAATAPVATSPNDISTSVYQSEVYVTEEGTAGV 1140
DB 1021 SHTGPTFPVVPVPPFSSITAGVPTTAAATAPVATSPNDISTSVYQSEVYVTEEGTAGV 1080
QY 1141 ATSTGVATSGGLPIPPVSESPVLSVVSSTIIPVVISITSPSLQVPTSTSIIVSSTA 1200
DB 1081 ATSTGVATSGGLPIPPVSESPVLSVVSSTIIPVVISITSPSLQVPTSTSIIVSSTA 1140

QY 1201 LPSVTSASASAGSSTATGPKPAPVAVSOQAAGSTTVGATLTSVSTTSFPPSTASQLS 1260
Db 1141 LPSVTSASASAGSSTATGPKPAPVAVSOQAAGSTTVGATLTSVSTTSFPPSTASQLS 1200
QY 1261 IQLSSSTSTPLTAETVAVSAHSLDKTSHSSTTGALAFSLSAPSSSSSPGAGVSSYSIQPGG 1330
Db 1201 IQLSSSTSTPLTAETVAVSAHSLDKTSHSSTTGALAFSLSAPSSSSSPGAGVSSYSIQPGG 1260
QY 1321 LHPVIVPVIASSTPLTPQAAGPTSTPLPOVPSIPPLVQVAVVAVQQLIHSQPQAL 1380
Db 1261 LHPVIVPVIASSTPLTPQAAGPTSTPLPOVPSIPPLVQVAVVAVQQLIHSQPQAL 1320
QY 1381 LHPVIVPVIASSTPLTPQAAGPTSTPLPOVPSIPPLVQVAVVAVQQLIHSQPQAL 1440
Db 1321 LHPVIVPVIASSTPLTPQAAGPTSTPLPOVPSIPPLVQVAVVAVQQLIHSQPQAL 1380
QY 1441 TVTGTGIPPTAVAPSKLTLTSTSTCLPPTNLPLGVVALPVTVPVTPGVSTPVSTTSQVK 1500
Db 1381 TVTGTGIPPTAVAPSKLTLTSTSTCLPPTNLPLGVVALPVTVPVTPGVSTPVSTTSQVK 1440
QY 1501 PGTAAPKPLTKAPVLPVGTLPAGTLPSBQLPPEPGLTLOSOQPLEDLAOLRLTSP 1560
Db 1441 PGTAAPKPLTKAPVLPVGTLPAGTLPSBQLPPEPGLTLOSOQPLEDLAOLRLTSP 1500
QY 1561 EXITVTSAVGVMAAPTAITEAGTOPQKGVYQVKEGVVLATSSGAGVFKMGKRFQVAVAA 1620
Db 1501 EXITVTSAVGVMAAPTAITEAGTOPQKGVYQVKEGVVLATSSGAGVFKMGKRFQVAVAA 1560
QY 1621 DGAKKKGKNEKEDAKSVHFESESTSESSVLSSESTLTKPEPNCITTPGISSVPEBSAH 1680
Db 1561 DGAKKKGKNEKEDAKSVHFESESTSESSVLSSESTLTKPEPNCITTPGISSVPEBSAH 1620
QY 1681 KTTSSEAKSDTGOCTKXGRFOVTTTANKVGRFYSKTEDKITDPRKKGPPVAPSPFMLEQ 1740
Db 1621 KTTSSEAKSDTGOCTKXGRFOVTTTANKVGRFYSKTEDKITDPRKKGPPVAPSPFMLEQ 1680
QY 1741 AVLEPAVIPKKEPELSEPSHNLGPPSDPEAFLSRVDVDSGSPHSPHOLSSKSLPSQNL 1800
Db 1681 AVLEPAVIPKKEPELSEPSHNLGPPSDPEAFLSRVDVDSGSPHSPHOLSSKSLPSQNL 1740
QY 1801 SOSLSNFPNSYMSDNESDIEDDLKLELRDLADKHLKEIQDLSQKHEIESELYTKLG 1860
Db 1741 SOSLSNFPNSYMSDNESDIEDDLKLELRDLADKHLKEIQDLSQKHEIESELYTKLG 1800
QY 1861 KVPAPVAVIIPPAAPISGRRRRTKSGSKSSRSSLSGNKSPOLSGNSAASVLAHQOT 1920
Db 1801 KVPAPVAVIIPPAAPISGRRRRTKSGSKSSRSSLSGNKSPOLSGNSAASVLAHQOT 1860
QY 1921 LHPGNIPESSGQNLQPLKSPSSDNLYSAFTSDGAI SVPSLSAPQOQTSSTTVGATV 1980
Db 1861 LHPGNIPESSGQNLQPLKSPSSDNLYSAFTSDGAI SVPSLSAPQOQTSSTTVGATV 1920
QY 1981 NSQAQAQOPAPMNTSRKGTFTDHLKLVDMWAPAMNLSGRGSKGMNTBGPMAKFS 2040
Db 1921 NSQAQAQOPAPMNTSRKGTFTDHLKLVDMWAPAMNLSGRGSKGMNTBGPMAKFS 1980
QY 2041 APQOLCISMTSNLGGASAPISAAASATSLGHTKSCPCPOQYGFPAITPFGAOMSGGAPAPQ 2100
Db 1981 APQOLCISMTSNLGGASAPISAAASATSLGHTKSCPCPOQYGFPAITPFGAOMSGGAPAPQ 2040
QY 2101 PLGQFQVAVTASLQNFNINSLQKISINPSPSNLRTT 2136
Db 2041 PLGQFQVAVTASLQNFNINSLQKISINPSPSNLRTT 2076

APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
FILE REFERENCE: LEX-0178-USA
CURRENT FILING DATE: 2001-05-14
PRIORITY FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 40
LENGTH: 2169
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(2169)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-40

Query Match 96.7%; Score 10451.5; DB 4; Length 2169;
Beet Local Similarity 95.7%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

61 MDKDSRGAATTTTTEHFRFRSVYICDSNATALELPGLPISLPQPSIPAAVPOGAPPEPH 120
1 MDKDSRGAATTTTTEHFRFRSVYICDSNATALELPGLPISLPQPSIPAAVPOGAPPEPH 60
121 REETVAVATSOVAQOPPAAPAPGQAVAPAPSTVBSSTSKDRPVASQPSIVSGKEPPP 180
61 REETVAVATSOVAQOPPAAPAPGQAVAPAPSTVBSSTSKDRPVASQPSIVSGKEPPP 120
181 ARSGSGGSAKEPQEHSSQOQDDIEELTAVAGMSNDRFLKPIETIGRGSFKTVYGLD 240
121 ARSGSGGSAKEPQEHSSQOQDDIEELTAVAGMSNDRFLKPIETIGRGSFKTVYGLD 180
241 TETTVAVAKELQDRKTKSEQRFEKFAEMLKGLQHPNIVRFYDSMESTYKGGKCVLV 300
181 TETTVAVAKELQDRKTKSEQRFEKFAEMLKGLQHPNIVRFYDSMESTYKGGKCVLV 240
301 TELMTSGTLTKYLRKFKVMKIKVLRNCRQILKGLQPLATRTPEIHRDLKCDNIFITGP 360
241 TELMTSGTLTKYLRKFKVMKIKVLRNCRQILKGLQPLATRTPEIHRDLKCDNIFITGP 300
361 TGSVKTGDLATLKASFAKSVYIGPEFMAPEVBEKYSVDVYAFGCMLEMATSEY 420
301 TGSVKTGDLATLKASFAKSVYIGPEFMAPEVBEKYSVDVYAFGCMLEMATSEY 360
421 PYSECQNAQIYRVTSGVPASFDKVAIPVEKEIIEGCIROMKDERYSIKDLNHAFFQ 480
361 PYSECQNAQIYRVTSGVPASFDKVAIPVEKEIIEGCIROMKDERYSIKDLNHAFFQ 420
481 EETGVAVELAEEDDEKIAIKMLRIEDIKLKGXKXNDNAIFSPFLERDVEDVAQEM 540
421 EETGVAVELAEEDDEKIAIKMLRIEDIKLKGXKXNDNAIFSPFLERDVEDVAQEM 480
541 VESGVYCEGHKTMARAKIKRVSLIKRKEQRLVBEQKKEOEBESLKQVEQSSASQ 600
481 VESGVYCEGHKTMARAKIKRVSLIKRKEQRLVBEQKKEOEBESLKQVEQSSASQ 540
601 TGIKOLPASASTGIPASTTSASVSTQVEPEPADHQOLQYQOAPSISVLSGTVDSGGG 660
541 TGIKOLPASASTGIPASTTSASVSTQVEPEPADHQOLQYQOAPSISVLSGTVDSGGG 600
661 SSVFTESRVSSQOTVSGSHQEAHSTGVPHIPISTVQASOPHGVYPPSSVAQOSQSG 720
601 SSVFTESRVSSQOTVSGSHQEAHSTGVPHIPISTVQASOPHGVYPPSSVAQOSQSG 660
721 QPSSSILTVSSSOPLOHPQOQOQIOOTAPPOQTVQYSQTSSEATTAAQVSGQAP 780
661 QPSSSILTVSSSOPLOHPQOQOQIOOTAPPOQTVQYSQTSSEATTAAQVSGQAP 720
781 QVLPVAVSAGKQ----- 791

```

Db      721 QVLPOVSAGKQGFPPRLPPQPGDSNIAPSSNVASCIHSTVLXPPMPTEVLATPGFPPT 780
Qy      792 -----STQGVQVAPAEVAV 867
Db      781 VVQPVYESNLLVPMGVGVQGVQVVSQPGGSLAQAPTTSSQQAULBSTQGVQVAPAEVAV 840
Qy      808 ACPQATQPTTLASSVDSAHSDVASGMSDGNENVPSSSGRHEGRTTKRHYKRSYRSRHE 867
Db      841 ACPQATQPTTLASSVDSAHSDVASGMSDGNENVPSSSGRHEGRTTKRHYKRSYRSRHE 900
Qy      868 KTSRPKRLINVSNGKDRIVECOLETHNRMTVTFKFDLDGDNBEIATIMWNNDFILAE 927
Db      901 KTSRPKRLINVSNGKDRIVECOLETHNRMTVTFKFDLDGDNBEIATIMWNNDFILAE 960
Qy      928 RESFVQVREIIRKADMLSEDSVVEPEGQGLSLQKXDYGFSQSKLEGEFKQIP 987
Db      961 RESFVQVREIIRKADMLSEDSVVEPEGQGLSLQKXDYGFSQSKLEGEFKQIP 1020
Qy      988 SSMPOQIGITSSLTQVHSGRRFIVSPVESRLRESKYFSEITDTVAASTAQSPGM 1047
Db      1021 SSMPOQIGITSSLTQVHSGRRFIVSPVESRLRESKYFSEITDTVAASTAQSPGM 1080
Qy      1048 LSHSASSLSLQOAFSEIRRAQMTGENTAPPNESHGCTFPVVPPLSLAGVPTTAAT 1107
Db      1081 LSHSASSLSLQOAFSEIRRAQMTGENTAPPNESHGCTFPVVPPLSLAGVPTTAAT 1140
Qy      1108 APVATSSPNDISTSVIOSEVTVTEEGTAGVNTSGVTSGLPFPVSESPVLSVV 1167
Db      1141 APVATSSPNDISTSVIOSEVTVTEEGTAGVNTSGVTSGLPFPVSESPVLSVV 1200
Qy      1168 SSITIPAVSISTSPLOVPTSEIIVSSTALPVSATVATASAGSSTATPGPKPPA 1227
Db      1201 SSITIPAVSISTSPLOVPTSEIIVSSTALPVSATVATASAGSSTATPGPKPPA 1260
Qy      1228 VVSQQAAGSTTVGATLTSVSTTSPSTASQLSTQLSSSTSTPTLAFTVVVAHSLDKTS 1287
Db      1261 VVSQQAAGSTTVGATLTSVSTTSPSTASQLSTQLSSSTSTPTLAFTVVVAHSLDKTS 1320
Qy      1288 HSSSTGLAFSLASASSSSPGAGSVSYISQPGHPIVTSVASTTILLOQAAGPTSTPL 1347
Db      1321 HSSSTGLAFSLASASSSSPGAGSVSYISQPGHPIVTSVASTTILLOQAAGPTSTPL 1380
Qy      1348 LPQVPSIPPLVQVPAVPAVQOULIHQOPALPNOPTHCPVEDSDTOPKAGIDIK 1407
Db      1381 LPQVPSIPPLVQVPAVPAVQOULIHQOPALPNOPTHCPVEDSDTOPKAGIDIK 1440
Qy      1408 TLEBKRLSPSEHSSSGAQAASVLETSVTEGIPPTTAVAPSKLLTSTTSTCLP 1467
Db      1441 TLEBKRLSPSEHSSSGAQAASVLETSVTEGIPPTTAVAPSKLLTSTTSTCLP 1500
Qy      1468 TNLPLGLVALPVTVPVTPGVSTVSTTSGVKRGTAPSKRPLTKAVLPLVGTLEPLGTL 1527
Db      1501 TNLPLGLVALPVTVPVTPGVSTVSTTSGVKRGTAPSKRPLTKAVLPLVGTLEPLGTL 1560
Qy      1528 PSEOLPFPFPGSLTQSOQPLEDDAOLRRTISPEXITVTSAGVGVMAAPTAITEAGTOP 1587
Db      1561 PSEOLPFPFPGSLTQSOQPLEDDAOLRRTISPEXITVTSAGVGVMAAPTAITEAGTOP 1620
Qy      1588 QKGVSYQVKEGVLATSSGAGVFKMGRFQVSVAAQAKGEGKNKSEDAKSVFESSSTSESS 1647
Db      1621 QKGVSYQVKEGVLATSSGAGVFKMGRFQVSVAAQAKGEGKNKSEDAKSVFESSSTSESS 1680
Qy      1648 VLSSSSPBSTLVKREPCITIGISDVPEBAHKTSEAKSDPGQPKVCRQVTTAN 1707
Db      1681 VLSSSSPBSTLVKREPCITIGISDVPEBAHKTSEAKSDPGQPKVCRQVTTAN 1740
Qy      1708 KVGFSSSKTEBDKLTDTKKKGVPVAPPPMDLEQAVLPAVLPKKEKPELSEBSHNGRPSD 1767
Db      1741 KVGFSSSKTEBDKLTDTKKKGVPVAPPPMDLEQAVLPAVLPKKEKPELSEBSHNGRPSD 1800
Qy      1768 PEAAPLSBDVDSGSPHSFHQSSKSLPSQNLQSLSNFSNSSYMSDNESEDIEDLX 1827

```

```

Db      1801 PEAAPLSBDVDSGSPHSFHQSSKSLPSQNLQSLSNFSNSSYMSDNESEDIEDLX 1860
Qy      1828 LELRLRDLKHEIQLDLSRQKHEIESLYTKLGKVPVAVIIPPAADLSGRRRPTSKGS 1887
Db      1861 LELRLRDLKHEIQLDLSRQKHEIESLYTKLGKVPVAVIIPPAADLSGRRRPTSKGS 1920
Qy      1888 KSSRSSSLGNKSPQLSGNLSGOSASVLPQOULHPPGNI PESGQOULLPLKPSPSDN 1947
Db      1921 KSSRSSSLGNKSPQLSGNLSGOSASVLPQOULHPPGNI PESGQOULLPLKPSPSDN 1980
Qy      1948 LYSAPTSDAIIVPSLSAPCGQSTSTNTGATNSQAQAQAPAMTSSRGTEETDHLK 2007
Db      1981 LYSAPTSDAIIVPSLSAPCGQSTSTNTGATNSQAQAQAPAMTSSRGTEETDHLK 2040
Qy      2008 VDNWADAMNLSGRSGSKGHNVEGPMARKFSAPOLCISMTSNLGSAPISAASATSL 2067
Db      2041 VDNWADAMNLSGRSGSKGHNVEGPMARKFSAPOLCISMTSNLGSAPISAASATSL 2100
Qy      2068 GHFTKSCMPQOYGFPATPFGAQWSGTGAPQPLQOFPVGTASLQNFNIMLQKISN 2127
Db      2101 GHFTKSCMPQOYGFPATPFGAQWSGTGAPQPLQOFPVGTASLQNFNIMLQKISN 2160
Qy      2128 PGGSNLRRT 2136
Db      2161 PGGSNLRRT 2169

RESULT 9
US-09-854-856-34
; Sequence 34, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hillman, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2001-05-14
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 2322
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)-(2322)
; OTHER INFORMATION: xaa = Any Amino Acid
US-09-854-856-34

Query Match          96.0%; Score 10375; DB 4; Length 2322;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 0; Indels 246; Gaps 1;

```

Db 181 TETVEAMCELODRKLTSEORKEBEAMLKGLQHPNIVRFYDWSWETVKGKCCIVLV 240
 QY 301 TELMTSGTLKTVLKRFKMKTKVLSMGRQILKGIQFHTTTPPIIYHDLKCDNIFITGP 360
 Db 241 TELMTSGTLKTVLKRFKMKTKVLSMGRQILKGIQFHTTTPPIIYHDLKCDNIFITGP 300
 QY 361 TGSVXIGLGLATLKRASFASVIGTPEPFMAPEMYEEXYDESVDVYAGMCMLEMA7SEY 420
 Db 301 TGSVXIGLGLATLKRASFASVIGTPEPFMAPEMYEEXYDESVDVYAGMCMLEMA7SEY 360
 QY 421 PYSECQNAAOIYRYRYTSVGVKASPDKVAIPVKEIIEGCIYRQNDERYSIKDLNHAFFQ 480
 Db 361 PYSECQNAAOIYRYRYTSVGVKASPDKVAIPVKEIIEGCIYRQNDERYSIKDLNHAFFQ 420
 QY 481 EETGVRYLAEDDEGEKIAIKLMLRIEDIKLKKGKYNENAIERSFDLERVPEVDVACEM 540
 Db 421 EETGVRYLAEDDEGEKIAIKLMLRIEDIKLKKGKYNENAIERSFDLERVPEVDVACEM 480
 QY 541 VESGYVEGDHKTMAKIKORVSLIKRREOROLVREOEKKOBEESLKQOEVASASQ 600
 Db 481 VESGYVEGDHKTMAKIKORVSLIKRREOROLVREOEKKOBEESLKQOEVASASQ 540
 QY 601 TGIQOLASASTGITASTASVSTOVEPEPEADQHOQLOQOPSLISVSDGTVDSGQG 660
 Db 541 TGIQOLASASTGITASTASVSTOVEPEPEADQHOQLOQOPSLISVSDGTVDSGQG 600
 QY 661 SSVFTESRVSQOQTVSYGSOHEOASHGTVPNGHPSYWOQASOPHGVYPPESVAAQGSQ 720
 Db 601 SSVFTESRVSQOQTVSYGSOHEOASHGTVPNGHPSYWOQASOPHGVYPPESVAAQGSQ 660
 QY 721 QPSSSLTGVSSSQPIQHPQOQOQGIQOTAPPOQTVYSLSQTSSTSEATTAQVPSOPQAP 780
 Db 661 QPSSSLTGVSSSQPIQHPQOQOQGIQOTAPPOQTVYSLSQTSSTSEATTAQVPSOPQAP 720
 QY 781 QVLPVSAAGQ----- 791
 Db 721 QVLPVSAAGQVPSQPEVPTIOGEPQIPVATQPSVPEVHSGAHPLPVGQPLPPLPQY 780
 QY 792 ----- 791
 Db 781 VSOQIPSTPHVSTAQTFSSLPITMAAGITQPLITLASATTAIPGVSTVVPBQPLPPL 840
 QY 792 ----- 791
 Db 841 QPVTOLPSQVHPOLLQPAVQSMGIPANLGOAABVPLSSGDVLYOGFPPLPPOVPGDSNT 900
 QY 792 ----- 791
 Db 901 APSSNVAVCIHSTVLAPMPTEVLATPGYFPVTVQPYVESNLLVPMGAGVQGVQVQSQ 960
 QY 792 -----STQVGSQVAPAEVAVANOPATQPTTLASVDSAHSDVAGMS 834
 Db 961 GSILAQAPTTSSQOAVLESTQVGSQVAPAEVAVANOPATQPTTLASVDSAHSDVAGMS 1020
 QY 835 DGENEVSSSSGRHGRITTKRHYKVSRSRHEKTSRPKLILNVSNKGRVVECOLETH 894
 Db 1021 DGENEVSSSSGRHGRITTKRHYKVSRSRHEKTSRPKLILNVSNKGRVVECOLETH 894
 QY 895 NRGKVPFKPLDNDNPELITIMVNDFLAIERESVVDVREIIEKADMLSEBVSVER 954
 Db 1081 NRGKVPFKPLDNDNPELITIMVNDFLAIERESVVDVREIIEKADMLSEBVSVER 954
 QY 955 EGDQGESLQGDYDYGFSQOKLEGEFKOPIPASWMOQIGIPTSLTQVYHSGRRFTY 1014
 Db 1141 EGDQGESLQGDYDYGFSQOKLEGEFKOPIPASWMOQIGIPTSLTQVYHSGRRFTY 1014
 QY 1015 SPVPSRLRBSKVPESEITDTVAASTAQSPAMNLSHASSLSLQAFSELRRAQOMTEGPN 1074
 Db 1201 SPVPSRLRBSKVPESEITDTVAASTAQSPAMNLSHASSLSLQAFSELRRAQOMTEGPN 1074
 QY 1075 TAPPNSSHGPTFPVVPPIPLSIAQVPTTAATAFVPAISPPNDISTSVTQSEVTPTE 1134
 Db 1261 TAPPNSSHGPTFPVVPPIPLSIAQVPTTAATAFVPAISPPNDISTSVTQSEVTPTE 1134

QY 1135 EGIAGVNTSGVNTSGGLPIPPVSESPVSSVSSITIPAVVSIITSPSLOVPTSEI 1194
 Db 1321 EGIAGVNTSGVNTSGGLPIPPVSESPVSSVSSITIPAVVSIITSPSLOVPTSEI 1380
 QY 1195 VVSSTALYPSTVATASAGSTATGPKRPVAVSQAAASTVVGATLTSVSTTSPS 1254
 Db 1381 VVSSTALYPSTVATASAGSTATGPKRPVAVSQAAASTVVGATLTSVSTTSPS 1440
 QY 1255 TASQLSIOLSSSTPTLAEVTVVSAHSLDKTSHSTTGLAFSTISAPSSSSPGAGVSY 1314
 Db 1441 TASQLSIOLSSSTPTLAEVTVVSAHSLDKTSHSTTGLAFSTISAPSSSSPGAGVSY 1500
 QY 1315 ISOPGGLHPLVIEVNTASTPLLPQAGPTPLLPVPSIPPLVOPVAVVQOQLIHS 1374
 Db 1501 ISOPGGLHPLVIEVNTASTPLLPQAGPTPLLPVPSIPPLVOPVAVVQOQLIHS 1560
 QY 1375 QOPQALLPNOPTHCPVSDPTOPKAPGIDDKTLERKLSLFEHSSSGAQHASVLET 1434
 Db 1561 QOPQALLPNOPTHCPVSDPTOPKAPGIDDKTLERKLSLFEHSSSGAQHASVLET 1620
 QY 1435 SLVIESTVTGPIPTTAVAPSKLITSTSTCLPPTNLPLGVALPVTVPVTEGQVSTEVST 1494
 Db 1621 SLVIESTVTGPIPTTAVAPSKLITSTSTCLPPTNLPLGVALPVTVPVTEGQVSTEVST 1680
 QY 1495 TTSGVKRGTAPSKPBLTKAVLPVGTLPAGTLPSEQLPPPPGSLTQSQOPLDLDLQOL 1554
 Db 1681 TTSGVKRGTAPSKPBLTKAVLPVGTLPAGTLPSEQLPPPPGSLTQSQOPLDLDLQOL 1740
 QY 1555 RRTLSPEXITVTSVAVGVMAAPALTEAGTOPQKGVSOVKEGVLATSSAGVFKMGRF 1614
 Db 1741 RRTLSPEXITVTSVAVGVMAAPALTEAGTOPQKGVSOVKEGVLATSSAGVFKMGRF 1800
 QY 1615 QVSVADGAQKEGKNSEDAKSVHESSTSESSVSLSSSSESTLYKEPNGITTPGISSD 1674
 Db 1801 QVSVADGAQKEGKNSEDAKSVHESSTSESSVSLSSSSESTLYKEPNGITTPGISSD 1860
 QY 1675 VPESAKHTTSAEAKSDTGOPTKVGRFQVYTTTANKVGFVSVKTEBKTDTJTKKGPVAPSP 1734
 Db 1861 VPESAKHTTSAEAKSDTGOPTKVGRFQVYTTTANKVGFVSVKTEBKTDTJTKKGPVAPSP 1920
 QY 1735 FMDLEQAVLPVAVIPKKEKPELSEPSHNGSPSPDEAFLSRVDVDDGSGSPHSHOLSSKS 1794
 Db 1921 FMDLEQAVLPVAVIPKKEKPELSEPSHNGSPSPDEAFLSRVDVDDGSGSPHSHOLSSKS 1980
 QY 1795 LPSQNLSSQSLSNSFNSSYMSDNESEDIEDLKLIEDRLRDLKXELQIDLQSHQKHIES 1854
 Db 1981 LPSQNLSSQSLSNSFNSSYMSDNESEDIEDLKLIEDRLRDLKXELQIDLQSHQKHIES 2040
 QY 1855 LYTKLGKVPVPAVILPPAAPLSGRRRPTYSKGSKSRSSSLGNKSPOLSGNLGQSAASV 1914
 Db 2041 LYTKLGKVPVPAVILPPAAPLSGRRRPTYSKGSKSRSSSLGNKSPOLSGNLGQSAASV 2100
 QY 1915 LHPQOQLHPPGNIPESGQOQLQPLKPSPESDMLYGAFTSDGASVYSLSAPFGGTSSTN 1974
 Db 2101 LHPQOQLHPPGNIPESGQOQLQPLKPSPESDMLYGAFTSDGASVYSLSAPFGGTSSTN 2160
 QY 1975 TVGATVNSQAQAQOPPAMTSSRKGTFTDDIHLKLVDMWARDAMNLSGRGSGKGMNTEGPG 2034
 Db 2161 TVGATVNSQAQAQOPPAMTSSRKGTFTDDIHLKLVDMWARDAMNLSGRGSGKGMNTEGPG 2220
 QY 2035 MARKFSAPQOLCISMTSNLGGSPISAAASVTSIGHFTTKSCBPQOYGFAPATPGAQMSGT 2094
 Db 2221 MARKFSAPQOLCISMTSNLGGSPISAAASVTSIGHFTTKSCBPQOYGFAPATPGAQMSGT 2280
 QY 2095 GGFAPQPLQOPQVGTASLQNFNLSNLOKSIINPPPSNLRKT 2136
 Db 2281 GGFAPQPLQOPQVGTASLQNFNLSNLOKSIINPPPSNLRKT 2322

RESULT 10
 US-09-854-856-62
 ; Sequence 62, Application us/09854856

Patent No. 6541252
 GENERAL INFORMATION:
 APPLICANT: Walke, D. Made
 APPLICANT: Hilbun, Erin
 APPLICANT: Donoho, Gregory
 APPLICANT: Turner, C. Alexander Jr.
 TITLE OF INVENTION: No. 6541252 Human Kinases and Polynucleotides
 FILE REFERENCE: LEX-0178-USA
 CURRENT APPLICATION NUMBER: US/09/854,856
 CURRENT FILING DATE: 2001-05-14
 PRIOR APPLICATION NUMBER: US 60/206,015
 PRIOR FILING DATE: 2000-05-19
 NUMBER OF SEQ ID NOS: 64
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 62
 LENGTH: 2048
 TYPE: PR
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: VARIANT
 LOCATION: (1)...(2048)
 OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-854-856-62

Query Match 95.7%; Score 10344; DB 4; Length 2048;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 2048; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

61 MDKDSGMAATTTTTEHFRFRSVICDSNATLELPGLPLSLPQSPISPAVPOGAPPEPH 120
 1 MDKDSGMAATTTTTEHFRFRSVICDSNATLELPGLPLSLPQSPISPAVPOGAPPEPH 60
 121 REEVNTATISQVAAQOPPAAPAEQAVAGAPSTVSSSTSKRPVOPSLVSGKEPP 180
 61 REEVNTATISQVAAQOPPAAPAEQAVAGAPSTVSSSTSKRPVOPSLVSGKEPP 120
 181 ARSGSGGSAKEPEERSQODDIEELKAVGMSNGRFLKFDIEIGRSGFKTVYGLD 240
 121 ARSGSGGSAKEPEERSQODDIEELKAVGMSNGRFLKFDIEIGRSGFKTVYGLD 180
 241 TETVEVAMCELODRUKTKSERORFKEAEMLKGLQHPNIVRFYDSESTVKGKCIYLV 300
 181 TETVEVAMCELODRUKTKSERORFKEAEMLKGLQHPNIVRFYDSESTVKGKCIYLV 240
 301 TELMTSGTLKTYLKRFRVMKTVLRSMCRQILKGLQPLHTPTPIIRDLKCDNIFITGP 360
 241 TELMTSGTLKTYLKRFRVMKTVLRSMCRQILKGLQPLHTPTPIIRDLKCDNIFITGP 300
 361 TGSVKGIGLGLATLKRPASFAKSVIGTPEFMAPEMYEEKYDESVDVYAFGMCLMATSEY 420
 301 TGSVKGIGLGLATLKRPASFAKSVIGTPEFMAPEMYEEKYDESVDVYAFGMCLMATSEY 360
 421 PYSECQNAQAQYREVTSGVCPASPDKVAIPEVKEIIIEGCIJRONDERYSIKDILNHAFFQ 480
 361 PYSECQNAQAQYREVTSGVCPASPDKVAIPEVKEIIIEGCIJRONDERYSIKDILNHAFFQ 420
 481 BETGVAVLAEEDDEKTAIKLMLRIEDIKLKGYKXNEAIEFSFDERIVPEDVAQEM 540
 421 BETGVAVLAEEDDEKTAIKLMLRIEDIKLKGYKXNEAIEFSFDERIVPEDVAQEM 480
 541 VESGVGCGDHKTAKAIKDRVSLIKRREOROLVREBQEKKKOEESLKKQVEOSSASQ 600
 481 VESGVGCGDHKTAKAIKDRVSLIKRREOROLVREBQEKKKOEESLKKQVEOSSASQ 540
 601 TGIRQLPSASTGIPASTTSASVSTOVEPEEPADDOHQLOYOQPSISVLSDGTVDSCG 660
 541 TGIRQLPSASTGIPASTTSASVSTOVEPEEPADDOHQLOYOQPSISVLSDGTVDSCG 600
 661 SSVTFESVSSQQTYSYSGHBOAHSTGTVEGHIPSTYQAOOSOPHGVPPSSVAGQSGG 720
 601 SSVTFESVSSQQTYSYSGHBOAHSTGTVEGHIPSTYQAOOSOPHGVPPSSVAGQSGG 660

721 QPSSSLTGVSSSQPIQHPQOQOIGIOQTAPPOQTVYISLSQTSSTSEATTAPQVSOQAP 780
 654 -----OQGIQOTAPPOQTVYISLSQTSSTSEATTAPQVSOQAP 692
 781 QVLPQVSAKQSTQGVQVAPPAVPAQAPATOPPTLASSVVSASHSDVAGMSDGENV 840
 693 QVLPQVSAKQSTQGVQVAPPAVPAQAPATOPPTLASSVVSASHSDVAGMSDGENV 752
 841 PSSSGHEGRTTTRHRYKSVRSRHEKTSRPLRLILANSNKDRVVECOLETHNRKMT 900
 753 PSSSGHEGRTTTRHRYKSVRSRHEKTSRPLRLILANSNKDRVVECOLETHNRKMT 812
 901 FKFDLDGNDPEEITATMVNNDFTLAIERESFVDQVEIIEKADMLSEDSVPEEGDGL 960
 813 FKFDLDGNDPEEITATMVNNDFTLAIERESFVDQVEIIEKADMLSEDSVPEEGDGL 872
 961 ESIQKDDYGFSSQKLEGFQKPIPASSMPOQIGIPTSLLQVHSAGRFTVSPVES 1020
 873 ESIQKDDYGFSSQKLEGFQKPIPASSMPOQIGIPTSLLQVHSAGRFTVSPVES 932
 1021 RLRESKVPSEITDTVAASTAQSPGNLSHASSLSLQCAFSELRRAQMTGENTAPNF 1080
 933 RLRESKVPSEITDTVAASTAQSPGNLSHASSLSLQCAFSELRRAQMTGENTAPNF 992
 1081 SHTGPTFPVVPPLSSIAQVPTTAATAPVATSPSPNDISTSVIOSEVTPPEEGTAGV 1140
 993 SHTGPTFPVVPPLSSIAQVPTTAATAPVATSPSPNDISTSVIOSEVTPPEEGTAGV 1052
 1141 ATSGVATSGGLPIPPVPSBPVLSVSSITTPAVVISITTSLSQVPTSEIIVSSTA 1200
 1053 ATSGVATSGGLPIPPVPSBPVLSVSSITTPAVVISITTSLSQVPTSEIIVSSTA 1112
 1201 LYPSTVATSAAGSGTATPGKPPAVVSOQAAGSTTVGATLTSVTTSPSTASQLS 1260
 1113 LYPSTVATSAAGSGTATPGKPPAVVSOQAAGSTTVGATLTSVTTSPSTASQLS 1172
 1261 IQLSSSTPTLAEVTVVASHSLDKTSHSTTGLAFSLAPSSSSSPGAGVSYISOPGG 1320
 1173 IQLSSSTPTLAEVTVVASHSLDKTSHSTTGLAFSLAPSSSSSPGAGVSYISOPGG 1232
 1321 LHPVLPVSVASTPILPQAAGPTSTPLPQVPSIPPLVQVAVNPAVQOQLIHQOPAL 1380
 1233 LHPVLPVSVASTPILPQAAGPTSTPLPQVPSIPPLVQVAVNPAVQOQLIHQOPAL 1292
 1381 LHPVLPVSVASTPILPQAAGPTSTPLPQVPSIPPLVQVAVNPAVQOQLIHQOPAL 1440
 1293 LHPVLPVSVASTPILPQAAGPTSTPLPQVPSIPPLVQVAVNPAVQOQLIHQOPAL 1352
 1441 TVTPGIPPTAVAPSKLLTSTSTCLPPTNLPLGTVALPVTVPVPGQVSTPVSTTSGVK 1500
 1353 TVTPGIPPTAVAPSKLLTSTSTCLPPTNLPLGTVALPVTVPVPGQVSTPVSTTSGVK 1412
 1501 PGRAPSRPILTKAPVLPVGELEPAGTLPSBOLPFPBPSPILQSOQPLEDDAOLRLTSP 1560
 1413 PGRAPSRPILTKAPVLPVGELEPAGTLPSBOLPFPBPSPILQSOQPLEDDAOLRLTSP 1472
 1561 EXITVTGAVPVSAAATTAITEAGTOPQKGVSYQYKEGPVLATSSGAVFPMGRPOVAA 1620
 1473 EXITVTGAVPVSAAATTAITEAGTOPQKGVSYQYKEGPVLATSSGAVFPMGRPOVAA 1532
 1621 DGAQKGNKSEDAKSVHPESTSESSVLSSSSESTLVKPEPNCITIPGISVPEPSAH 1680
 1533 DGAQKGNKSEDAKSVHPESTSESSVLSSSSESTLVKPEPNCITIPGISVPEPSAH 1592
 1681 KTTASEAKSDTQGTQYGRFOVTTTAKVGFVSFKTEKTLTDPKKGPVAPPFMDLEQ 1740
 1593 KTTASEAKSDTQGTQYGRFOVTTTAKVGFVSFKTEKTLTDPKKGPVAPPFMDLEQ 1652
 1741 AVLPVAVIPKKEKPELSEPSHLNGSPSPEAFLSRVDVDSGSPHSBHQSLSPQNTL 1800
 1653 AVLPVAVIPKKEKPELSEPSHLNGSPSPEAFLSRVDVDSGSPHSBHQSLSPQNTL 1712
 1801 SGLSNSFNSSYMSDNESEDIEDDLKLELRRLDKHLKEIQDQSKRKHIEBLTYTKG 1860

Db 1713 SQSLSNSFNSSWSSDSDNDEDEDLKLELRRLRDKHKEIODLSQKHEIESLYTKLG 1772

Qy 1861 KVPAAVILPPAPALSGRRRPTKSGKSSRSSLSGNKSPOLSGLSQSAASVHHPOT 1920

Db 1773 KVPAAVILPPAPALSGRRRPTKSGKSSRSSLSGNKSPOLSGLSQSAASVHHPOT 1832

Qy 1921 LHPNPIESGONQLLPKSPSSSDNTYSAFTSDGALSVPSLAPAGQGSSTNTVATV 1980

Db 1833 LHPNPIESGONQLLPKSPSSSDNTYSAFTSDGALSVPSLAPAGQGSSTNTVATV 1892

Qy 1981 NSQAQAQPPAMTSRKGTFTDHLKLVDMNARDAMNLSGRGSGKHNNYEGPGMARFES 2040

Db 1893 NSQAQAQPPAMTSRKGTFTDHLKLVDMNARDAMNLSGRGSGKHNNYEGPGMARFES 1952

Qy 2041 APGOLCISMTSNLGSAPISASATSLSGFTKSMCPPOQYGPATPGAGWSGTGGAPQ 2100

Db 1953 APGOLCISMTSNLGSAPISASATSLSGFTKSMCPPOQYGPATPGAGWSGTGGAPQ 2012

Qy 2101 PLGQFQPVGTASLQNFNINLQKSIISNPPGSLRTT 2136

Db 2013 PLGQFQPVGTASLQNFNINLQKSIISNPPGSLRTT 2048

RESULT 11

US-09-854-856-56
 ; Sequence 56, Application US/09854856
 ; Patent No. 6541252
 ; GENERAL INFORMATION:
 ; APPLICANT: Walke, D. Wade
 ; APPLICANT: Hilbun, Erin
 ; APPLICANT: Donoho, Gregory
 ; APPLICANT: Turner, C. Alexander Jr.
 ; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
 ; FILE REFERENCE: LEX-0178-USA
 ; CURRENT APPLICATION NUMBER: US/09/854,856
 ; PRIORITY FILING DATE: 2001-05-14
 ; PRIOR APPLICATION NUMBER: US 60/206,015
 ; NUMBER OF SEQ ID NOS: 64
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 56
 ; LENGTH: 2141
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(2141)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 ; US-09-854-856-56

Query Match 95.1%; Score 10287.5; DB 4; Length 2141;
 Best Local Similarity 94.4%; Pred. No. 0;
 Matches 2048; Conservative 0; Mismatches 0; Indels 121; Gaps 2;

Qy 61 MDKDSGAAAATTTTTEHRRFRFRSVICDSNATALELDELPLSLPQPSIPAAVPOSAPEPH 120

Db 1 MDKDSGAAAATTTTTEHRRFRFRSVICDSNATALELDELPLSLPQPSIPAAVPOSAPEPH 60

Qy 121 RETTVATATISQVAQOPPAAPAAAGVAVAPAPSTVPSSTSKORPVASOPSLVSGKEPP 180

Db 61 RETTVATATISQVAQOPPAAPAAAGVAVAPAPSTVPSSTSKORPVASOPSLVSGKEPP 120

Qy 181 ASGSGSGSAKEPQERSQOQDDIELETKAVMSNDGRFLKFDIEIGSGFETVYKGLD 240

Db 121 ASGSGSGSAKEPQERSQOQDDIELETKAVMSNDGRFLKFDIEIGSGFETVYKGLD 180

Qy 241 TETTVAVANCELQDRKLTKSERQRFKEAEMLKGLQHPNIVRFYDSWESTVKGKCIIV 300

Db 181 TETTVAVANCELQDRKLTKSERQRFKEAEMLKGLQHPNIVRFYDSWESTVKGKCIIV 240

Qy 301 TELMSTGLTKYTKRKWKIKVLRSWCROILKGLQFLHTTRPPIIHRDLKCDNIFITGP 360

Db 241 TELMSTGLTKYTKRKWKIKVLRSWCROILKGLQFLHTTRPPIIHRDLKCDNIFITGP 300

Qy 361 TGSYKIGDLGATIKRASFAKSVIGTPEFMAPEWEEKYKSDSDVYAFAGMCLMATASEY 420

Db 301 TGSYKIGDLGATIKRASFAKSVIGTPEFMAPEWEEKYKSDSDVYAFAGMCLMATASEY 360

Qy 421 PYSCQNAAQYRRVTSQVKASFDKVAIPVEKTEIECCIQONKDERISYDOLNHAFO 480

Db 361 PYSCQNAAQYRRVTSQVKASFDKVAIPVEKTEIECCIQONKDERISYDOLNHAFO 420

Qy 481 EETGVARELAEDDEGEKIALKMLRIEDIKLKGKYKQNEAIEPSFDERDVEDVAQEM 540

Db 421 EETGVARELAEDDEGEKIALKMLRIEDIKLKGKYKQNEAIEPSFDERDVEDVAQEM 480

Qy 541 VESGYVEBGDKHTMAKAIKDRVSLIKRRREGQRLVREEQKKOEBSLKQOVEGSSAQ 600

Db 481 VESGYVEBGDKHTMAKAIKDRVSLIKRRREGQRLVREEQKKOEBSLKQOVEGSSAQ 540

Qy 601 TGIKQLPASASTGIPATASVSTOVEPEPEADQHQLOQYQOPSSISVSDGTVDGQG 660

Db 541 TGIKQLPASASTGIPATASVSTOVEPEPEADQHQLOQYQOPSSISVSDGTVDGQG 600

Qy 661 SSYFTESRVSSQQTVSYSQHEQAHSTGTVGHIPTVOAOSOPHGYPPSSVAQGOQSG 720

Db 601 SSYFTESRVSSQQTVSYSQHEQAHSTGTVGHIPTVOAOSOPHGYPPSSVAQGOQSG 660

Qy 721 QPSSSLITGVSSSOPICHPQOQCGIQQTPAPQOTVOYSLQSTSSBATTAAQVSPQAP 780

Db 654 QPSSSLITGVSSSOPICHPQOQCGIQQTPAPQOTVOYSLQSTSSBATTAAQVSPQAP 692

Qy 781 QVLPOVSAGKQ-----QOGIQQTPAPQOTVOYSLQSTSSBATTAAQVSPQAP 791

Db 693 QVLPOVSAGKQPPPLPPYPPGDSNAPSNNASVCISHTVLXPPMPTEVLATPGYPT 752

Qy 792 -----STQGSQVAPAEVAV 807

Db 753 VVQPYEENLVPVGGVGVGVQVQSPGSLAQAPPTSSQQAVALSTQGVQVAPAEVAV 812

Qy 808 AQPQATOPTTLASVDSASDVASGMSDGENVPSSSGRHEGRTTKHYKVSRSRHE 867

Db 813 AQPQATOPTTLASVDSASDVASGMSDGENVPSSSGRHEGRTTKHYKVSRSRHE 872

Qy 868 KTSRPKRLITLVNKNKDRVVECOLETHNRKMTVEKPELDGDNBEELATIMWNNDFILAE 927

Db 873 KTSRPKRLITLVNKNKDRVVECOLETHNRKMTVEKPELDGDNBEELATIMWNNDFILAE 932

Qy 928 RESFVQVREILIEKADEMLSEDSVPEPQDGLSELOKQDYGFSGSKLEGEFKQIP 987

Db 933 RESFVQVREILIEKADEMLSEDSVPEPQDGLSELOKQDYGFSGSKLEGEFKQIP 992

Qy 988 SSMPOQIGIPTSSLTOVHSAGRRFTVSPVESRLRESKYFPSEITDTVAASIAQSDGN 1047

Db 993 SSMPOQIGIPTSSLTOVHSAGRRFTVSPVESRLRESKYFPSEITDTVAASIAQSDGN 1052

Qy 1048 LSHSASSLSIQQAFSELRAQMTGEPNTAPPNNSHNGPFPVVPPLSLASAGPTTAAAT 1107

Db 1053 LSHSASSLSIQQAFSELRAQMTGEPNTAPPNNSHNGPFPVVPPLSLASAGPTTAAAT 1112

Qy 1108 APVPATSPNDISTSVIOSEVTVPTREGIAGVATGVTGSGGLPIPVSESPVLSVV 1167

Db 1113 APVPATSPNDISTSVIOSEVTVPTREGIAGVATGVTGSGGLPIPVSESPVLSVV 1172

Qy 1168 SSTTTPAVVISITSSLSLOVPTSTSIIVVSTALYSVVATASAGGSTATPGKPPA 1227

Db 1173 SSTTTPAVVISITSSLSLOVPTSTSIIVVSTALYSVVATASAGGSTATPGKPPA 1232

Qy 1228 VVSQAAGSTYGAALTSTSTTSPSTASQSLQSSSTSPPLAEVTVVASHSLDKTS 1287

Db 1233 VVSQAAGSTYGAALTSTSTTSPSTASQSLQSSSTSPPLAEVTVVASHSLDKTS 1292

Qy 1288 HSGTGLAFSLASPSSSSGAGVSSYISQPGLHPLVPSVIASTPIIPQAGPSTPL 1347

1293 HSSSTTGLAFSLASPSSSSSSGAGVSSYISQPGGLHPLVBSVIASTPILPQAAQPTSTPL 1352
 1348 LPQVPSIPLVOPANVPANVQOQLIHSGOPALLPNOPTHCEVSDTOPKAGIDIDK 1407
 1353 LPQVPSIPLVOPANVPANVQOQLIHSGOPALLPNOPTHCEVSDTOPKAGIDIDK 1412
 1408 TLEBKSLFSEHSSSSGAQHASVSLFSLVLESTVGTGIPITAVAPSKLLTSTSTCLPP 1467
 1413 TLEBKSLFSEHSSSSGAQHASVSLFSLVLESTVGTGIPITAVAPSKLLTSTSTCLPP 1472
 1468 TNLPLGTVALLPVTPVTPVPGQVSTVSTTSGVKGCTAPSKPLTKAYLVPGTLEPAGTL 1527
 1473 TNLPLGTVALLPVTPVTPVPGQVSTVSTTSGVKGCTAPSKPLTKAYLVPGTLEPAGTL 1532
 1528 PSEQLPPEPGLSLQSQOPLDLDOLRRTLSPEXITVTSVGVSAAPPAITAEAGTOP 1587
 1533 PSEQLPPEPGLSLQSQOPLDLDOLRRTLSPEXITVTSVGVSAAPPAITAEAGTOP 1592
 1588 QKGVSYQKEGVLATSSGAGVFKMGKRFQVSVADGAQKEGKNSKSDAKSVHFEESTSESS 1647
 1593 QKGVSYQKEGVLATSSGAGVFKMGKRFQVSVADGAQKEGKNSKSDAKSVHFEESTSESS 1652
 1648 VLSSESESTLVKPEPNGITIPGISDVPSAHKTTASEAKSDGQPTKVGKRFQVTTAN 1707
 1653 VLSSESESTLVKPEPNGITIPGISDVPSAHKTTASEAKSDGQPTKVGKRFQVTTAN 1712
 1708 KVGFSYKTEKEDITDTRKKEGVASPPFMDLEQAVLPAVLPKKEPELSEPSHLNGSPSD 1767
 1713 KVGFSYKTEKEDITDTRKKEGVASPPFMDLEQAVLPAVLPKKEPELSEPSHLNGSPSD 1772
 1768 PEAFLSRDVEDGSGSPHSPHQLSSKSLPSQNLQSLSNSFNSSYMSDNESEDIEDDLK 1827
 1773 PEAFLSRDVEDGSGSPHSPHQLSSKSLPSQNLQSLSNSFNSSYMSDNESEDIEDDLK 1832
 1828 LEIRLRDRLKHEIQLDLSROKHEIESLYTKLGKVPYPAVLPAPAPLSGRRRRPTKSKGS 1887
 1833 LEIRLRDRLKHEIQLDLSROKHEIESLYTKLGKVPYPAVLPAPAPLSGRRRRPTKSKGS 1892
 1888 KSSRSSSTSGKNSPOLSGMLSGOSASAVLHPQOQLHPENIPESONOLQPLKSPSSDN 1947
 1893 KSSRSSSTSGKNSPOLSGMLSGOSASAVLHPQOQLHPENIPESONOLQPLKSPSSDN 1952
 1948 LYSAFSTDGALSVLSLAPAGQSTSTNTGATVNSQAQAQPPAMTSSRKGTFTDHLKL 2007
 1953 LYSAFSTDGALSVLSLAPAGQSTSTNTGATVNSQAQAQPPAMTSSRKGTFTDHLKL 2012
 2008 VDNARADAMNLSGRRSGKGMNVBPGMARKFSAPOQLCTIMTSLGGSAPISASATSL 2067
 2013 VDNARADAMNLSGRRSGKGMNVBPGMARKFSAPOQLCTIMTSLGGSAPISASATSL 2072
 2068 GHFTKSMGPPOOYGPATPFGAOWSGTGAPAPQPLGQFQVGTASLQFNISNLOKSTSN 2127
 2073 GHFTKSMGPPOOYGPATPFGAOWSGTGAPAPQPLGQFQVGTASLQFNISNLOKSTSN 2132
 2128 PPGSNLRTT 2136
 2133 PPGSNLRTT 2141

PRIOR APPLICATION NUMBER: US 60/206,015
 PRIOR FILING DATE: 2000-05-19
 NUMBER OF SEQ ID NOS: 64
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 50
 LENGTH: 2294
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: VARIANT
 LOCATION: (1)-(2294)
 OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-854-856-50

Query Match 94.4%; Score 10211; DB 4; Length 2294;
 Best Local Similarity 88.2%; Pred. No. 0;
 Matches 2049; Conservative 0; Mismatches 0; Indels 274; Gaps 2;

61 MDKDSRGAATTTTTEHFRFRRSVTCDSNATALELPGLPLSLPQPSIPAAVPOGAPPEPH 120
 1 MDKDSRGAATTTTTEHFRFRRSVTCDSNATALELPGLPLSLPQPSIPAAVPOGAPPEPH 60

121 REETVATATSOVAQOPPAAPGEQAVAPASTVSSSTKDRPVSOPLVSGKEEPP 180
 61 REETVATATSOVAQOPPAAPGEQAVAPASTVSSSTKDRPVSOPLVSGKEEPP 120

181 ARSGSGGSAKEPOEERSQOQDDIELETRAVGMSNDGRFLKPDIEIGRSFKTVXGLD 240
 121 ARSGSGGSAKEPOEERSQOQDDIELETRAVGMSNDGRFLKPDIEIGRSFKTVXGLD 180

241 TETTEVAMCELODRKLTKEERORKEAEMLKGLQHPNIVRYEYDSWESTYKGCCTLV 300
 181 TETTEVAMCELODRKLTKEERORKEAEMLKGLQHPNIVRYEYDSWESTYKGCCTLV 240

301 TELMTSGTLTYLKRFRVMKIKVLRSMCRQILKGLQHLHRTPIIHRDLKCDNIPTGP 360
 241 TELMTSGTLTYLKRFRVMKIKVLRSMCRQILKGLQHLHRTPIIHRDLKCDNIPTGP 300

361 TGSVKIDGLIATLKRAFPKSVIGTEPFNAPEMEYEKYESVDVYAFGMCMLEMATSEY 420
 301 TGSVKIDGLIATLKRAFPKSVIGTEPFNAPEMEYEKYESVDVYAFGMCMLEMATSEY 360

421 PYSECQNAQOYRRVTSVGVKPAFDKVAIPBEKIIIEGCTIRONDERYSIDLHNAFQ 480
 361 PYSECQNAQOYRRVTSVGVKPAFDKVAIPBEKIIIEGCTIRONDERYSIDLHNAFQ 420

481 EETGVARELAEEEDGEEKIALKMLRIDIKLKXKYKONNEAIEFSFDERVPEPDAQEM 540
 421 EETGVARELAEEEDGEEKIALKMLRIDIKLKXKYKONNEAIEFSFDERVPEPDAQEM 480

541 VESGYVEGDHKTAKAIKDRVSLIKRREORQLVREOEKKQOESSLKQOVEQSSAQ 600
 481 VESGYVEGDHKTAKAIKDRVSLIKRREORQLVREOEKKQOESSLKQOVEQSSAQ 540

601 TGIRKQPSASTGIFTASTTASVSTQVPEEPADQHOQLOOQOQPSISVLSDGTVDSCG 660
 541 TGIRKQPSASTGIFTASTTASVSTQVPEEPADQHOQLOOQOQPSISVLSDGTVDSCG 600

661 SSVFTESSVSSQOQVSSQOQVSSQOQVSSQOQVSSQOQVSSQOQVSSQOQVSSQOQ 720
 601 SSVFTESSVSSQOQVSSQOQVSSQOQVSSQOQVSSQOQVSSQOQVSSQOQVSSQOQ 660

721 QPSSSLLTGVSQOPITQHPQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 780
 661 QPSSSLLTGVSQOPITQHPQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 720

781 QVLQVSNAGQO-----QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 791
 661 QVLQVSNAGQO-----QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 780

792 ----- 791
 661 ----- 791

793 VSOIPSTPHVSTVQOTGFSPLPTMAAGITOPDLTLASSATTALIPGVSTVPSQPLILL 812

QY 792 ----- 791
 Db 813 QPVTQLPSQVHPLLQPAVQSMGITPANIGQAAEVLSSGDVLYGCFPPRLPEQVPGDSNT 872
 QY 792 ----- 791
 Db 873 APSSNVA5C1HSITVLXPMPTEVLATPGYPTVQVYVESNLVPMGQVGGQVQVSGP 932
 QY 792 ----- 834
 Db 933 GSIAOAPFTSSQQAVALSEITQGV5QVAPAPVAAPOATQPTTLASSYD5ASHSDV5A5MS 992
 QY 835 DGENEVPS55GRHREKTRKHYKRSVRSRSHREKTSRKRLILVNSNGDRVBEQOLETH 894
 Db 993 DGENEVPS55GRHREKTRKHYKRSVRSRSHREKTSRKRLILVNSNGDRVBEQOLETH 1052
 QY 895 NRKAVTFEFDLDGDNPEEIIATIMVNDPILAIERESFVDQVREIIEKADEMI5EDV5VEP 954
 Db 1053 NRKAVTFEFDLDGDNPEEIIATIMVNDPILAIERESFVDQVREIIEKADEMI5EDV5VEP 1112
 QY 955 EGDGLESLOGKDYGFSGSQKLEEFKQPIPASMPQOIGIPSSLTQVYVH5AGRRITV 1014
 Db 1113 EGDGLESLOGKDYGFSGSQKLEEFKQPIPASMPQOIGIPSSLTQVYVH5AGRRITV 1172
 QY 1015 SPVESRLRESKVPSEITDTVA5TAQSPGMNL5H5ASLSLQOAFSELRRAQMTGEPN 1074
 Db 1173 SPVESRLRESKVPSEITDTVA5TAQSPGMNL5H5ASLSLQOAFSELRRAQMTGEPN 1232
 QY 1075 TAPNF5HTGTFPVVPPFLSS1AGVPTTAATAVAPVATSSPPNDISTSVI5QSEVTVPE 1134
 Db 1233 TAPNF5HTGTFPVVPPFLSS1AGVPTTAATAVAPVATSSPPNDISTSVI5QSEVTVPE 1292
 QY 1135 EGIAGVATSTVNV5GGLPIPV5ESPL5SV5SITTPAV51STTSP5I5QVTS5EI 1194
 Db 1293 EGIAGVATSTVNV5GGLPIPV5ESPL5SV5SITTPAV51STTSP5I5QVTS5EI 1352
 QY 1195 VVS5TALPV5TV5AT5ASAG5TATPGKPAV5V5QAA5STVGA5TLT5V5T5T5FP5 1254
 Db 1353 VVS5TALPV5TV5AT5ASAG5TATPGKPAV5V5QAA5STVGA5TLT5V5T5T5FP5 1412
 QY 1255 TAOGL5QL5SS5T5PTLAEVTV5ASL5DKT5H5ST5T5GAF5SI5AP5SS5SPAGV5SY 1314
 Db 1413 TAOGL5QL5SS5T5PTLAEVTV5ASL5DKT5H5ST5T5GAF5SI5AP5SS5SPAGV5SY 1472
 QY 1315 ISOPGGLHPIV5P5VIA5T5P5L5PQ5AGT5P5L5PQ5P5I5P5L5QV5ANV5PAV5Q5TLI5H5 1374
 Db 1473 ISOPGGLHPIV5P5VIA5T5P5L5PQ5AGT5P5L5PQ5P5I5P5L5QV5ANV5PAV5Q5TLI5H5 1532
 QY 1375 QOPQALLP5N5P5HT5CP5EVD5DTOPKAP5ID5IKT5LEK5LSL5F5E5H5SS5GA5Q5H5V5LET 1434
 Db 1533 QOPQALLP5N5P5HT5CP5EVD5DTOPKAP5ID5IKT5LEK5LSL5F5E5H5SS5GA5Q5H5V5LET 1592
 QY 1435 SLVIESVTV5GIP5T5TA5P5SKL5T5T5T5T5L5P5P5NL5PL5G5T5VAL5P5TV5V5T5P5G5V5T5P5V5T 1494
 Db 1593 SLVIESVTV5GIP5T5TA5P5SKL5T5T5T5T5L5P5P5NL5PL5G5T5VAL5P5TV5V5T5P5G5V5T5P5V5T 1652
 QY 1495 TTS5GVK5GT5P5R5P5L5TK5AP5VL5PV5G5TEL5P5AG5TL5P5R5D5L5P5P5G5SL5T5Q5O5P5L5ED5DA5Q 1554
 Db 1653 TTS5GVK5GT5P5R5P5L5TK5AP5VL5PV5G5TEL5P5AG5TL5P5R5D5L5P5P5G5SL5T5Q5O5P5L5ED5DA5Q 1712
 QY 1555 RRTLS5PEXITV5T5AV5GV5MA5P5AL5I5E5A5CT5Q5K5G5V5K5E5G5P5V5LAT5S5G5A5V5F5M5K5R5 1614
 Db 1713 RRTLS5PEXITV5T5AV5GV5MA5P5AL5I5E5A5CT5Q5K5G5V5K5E5G5P5V5LAT5S5G5A5V5F5M5K5R5 1772
 QY 1615 QV5V5AAD5G5Q5K5E5K5N5S5ED5K5V5H5F5B5S5I5E5S5V5L5SS5P5E5ST5LV5K5P5E5N5G5I5T5P5G5I5SD 1674
 Db 1773 QV5V5AAD5G5Q5K5E5K5N5S5ED5K5V5H5F5B5S5I5E5S5V5L5SS5P5E5ST5LV5K5P5E5N5G5I5T5P5G5I5SD 1832
 QY 1675 VP5SA5K5TT5A5S5E5K5S5T5G5P5TV5K5G5F5OV5TT5ANK5G5F5S5K5E5D5IT5D5K5K5G5P5A5SP 1734
 Db 1833 VP5SA5K5TT5A5S5E5K5S5T5G5P5TV5K5G5F5OV5TT5ANK5G5F5S5K5E5D5IT5D5K5K5G5P5A5SP 1892

QY 1735 FMDLEQAVLPAVITPKKEPELSEPSHLNGPSSDPEAATLSPDVDDGSGSPHSPHOL5SK5 1794
 Db 1893 FMDLEQAVLPAVITPKKEPELSEPSHLNGPSSDPEAATLSPDVDDGSGSPHSPHOL5SK5 1952
 QY 1795 LPSQNL5Q5LSN5SN5F5N5S5Y5M5SDN5SD5IED5DL5K5EL5RL5R5K5H5KEI5D5L5G5R5Q5H5I5E5 1854
 Db 1953 LPSQNL5Q5LSN5SN5F5N5S5Y5M5SDN5SD5IED5DL5K5EL5RL5R5K5H5KEI5D5L5G5R5Q5H5I5E5 2012
 QY 1955 LYTKL5GV5PP5AVI5T5P5A5P5L5G5RR5R5P5T5K5S5K5S5R5SS5LG5N5K5P5Q5L5SG5A5AV 1914
 Db 2013 LYTKL5GV5PP5AVI5T5P5A5P5L5G5RR5R5P5T5K5S5K5S5R5SS5LG5N5K5P5Q5L5SG5A5AV 2072
 QY 1915 LHPQOTLHP5P5N5I5E5G5Q5N5L5Q5L5K5P5SS5D5N5Y5A5T5S5G5A5V5P5L5AP5G5T5STN 1974
 Db 2073 LHPQOTLHP5P5N5I5E5G5Q5N5L5Q5L5K5P5SS5D5N5Y5A5T5S5G5A5V5P5L5AP5G5T5STN 2132
 QY 1975 TVGATVNS5Q5A5Q5AP5P5M5T5SR5K5GT5FD5DL5H5K5LD5V5N5A5R5D5M5N5L5G5R5G5S5G5H5N5Y5E5G5 2034
 Db 2133 TVGATVNS5Q5A5Q5AP5P5M5T5SR5K5GT5FD5DL5H5K5LD5V5N5A5R5D5M5N5L5G5R5G5S5G5H5N5Y5E5G5 2192
 QY 2035 MARK5SAP5G5OL5C5I5M5T5N5L5G5S5A5P5I5S5A5S5T5L5G5H5FT5K5M5CP5Q5Y5G5P5AT5P5G5A5W5G5T 2094
 Db 2193 MARK5SAP5G5OL5C5I5M5T5N5L5G5S5A5P5I5S5A5S5T5L5G5H5FT5K5M5CP5Q5Y5G5P5AT5P5G5A5W5G5T 2252
 QY 2095 GGPAP5P5L5G5F5Q5P5V5G5T5A5L5Q5N5F5N5L5O5K5I5S5N5P5G5N5L5RT 2136
 Db 2253 GGPAP5P5L5G5F5Q5P5V5G5T5A5L5Q5N5F5N5L5O5K5I5S5N5P5G5N5L5RT 2294

RESULT 13
 US-09-854-856-16
 : Sequence 16, Application US/09854856
 : Patent No. 6541252
 : GENERAL INFORMATION:
 : APPLICANT: Walke, D. Wade
 : APPLICANT: Hilbun, Erin
 : APPLICANT: Donoho, Gregory
 : APPLICANT: Turner, C. Alexander Jr.
 : TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
 : TITLE OR INVENTION: Encoding the Same
 : FILE REFERENCE: LEX-0178-USA
 : CURRENT APPLICATION NUMBER: US/09/854,856
 : CURRENT FILING DATE: 2001-05-14
 : PRIOR APPLICATION NUMBER: US 60/206,015
 : PRIOR FILING DATE: 2000-05-19
 : NUMBER OF SEQ ID NOS: 64
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 16
 : LENGTH: 1999
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: VARIANT
 : LOCATION: (1)...(1999)
 : OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-854-856-16

Query Match 91.8%; Score 9929; DB 4; Length 1999;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1972; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

QY 1 MSGGA5EK5ST5G5SL5L5P5P5AP5P5K5N5G5S5D5S5V5G5E5K5L5G5A5A5D5A5V5G5T5E5Y5R5R5R5T 60
 Db 1 MSGGA5EK5ST5G5SL5L5P5P5AP5P5K5N5G5S5D5S5V5G5E5K5L5G5A5A5D5A5V5G5T5E5Y5R5R5R5T 60
 QY 61 MDKDS5GA5A5T5T5T5T5H5R5F5R5S5V5I5C5D5N5A5T5A5L5E5L5P5L5P5L5P5S5I5P5A5V5P5Q5A5P5E5P5H 120
 Db 61 MDKDS5GA5A5T5T5T5T5H5R5F5R5S5V5I5C5D5N5A5T5A5L5E5L5P5L5P5L5P5S5I5P5A5V5P5Q5A5P5E5P5H 120
 QY 121 REETV5A5T5S5V5A5Q5P5P5A5A5P5G5Q5A5V5A5G5P5A5T5V5P5S5T5K5D5R5P5V5Q5P5L5V5G5S5K5E5P5P 180
 Db 121 REETV5A5T5S5V5A5Q5P5P5A5A5P5G5Q5A5V5A5G5P5A5T5V5P5S5T5K5D5R5P5V5Q5P5L5V5G5S5K5E5P5P 180

QY 181 ARSGSGGSAKPEGEERSQOQDDIELEETAYAGKMSNDGRFLKEDIIEIGRGSFKTVYKGLD 240
 Db 181 ARSGSGGSAKPEGEERSQOQDDIELEETAYAGKMSNDGRFLKEDIIEIGRGSFKTVYKGLD 240
 QY 241 TETTVAVAKCELQDRKLTSEBQAFKAEMLKGLQHPNIVRPFDSWESTYKGGKCLVAV 300
 Db 241 TETTVAVAKCELQDRKLTSEBQAFKAEMLKGLQHPNIVRPFDSWESTYKGGKCLVAV 300
 QY 301 TELMTSGTLKTYLKRFFVMKIKVLRSMCRQILKGLQHLTRTPPIIHRDCKDNIFITGP 360
 Db 301 TELMTSGTLKTYLKRFFVMKIKVLRSMCRQILKGLQHLTRTPPIIHRDCKDNIFITGP 360
 QY 361 TGSVKIGDLGLATLKRAFAKSVLGTPEFMAPEMYEKYDESVDVYAFGMCMEMATSEY 420
 Db 361 TGSVKIGDLGLATLKRAFAKSVLGTPEFMAPEMYEKYDESVDVYAFGMCMEMATSEY 420
 QY 421 PYBRCQMAAQIYRRVTSGVVPASFDKVAIBVKIIEGCTRONDERYSIKDLINAFQ 480
 Db 421 PYBRCQMAAQIYRRVTSGVVPASFDKVAIBVKIIEGCTRONDERYSIKDLINAFQ 480
 QY 481 BETGVRELAEEDGEEKIAIKMLRIDIKKLKGYKONEAIFSPLEBDVPEDAQEM 540
 Db 481 BETGVRELAEEDGEEKIAIKMLRIDIKKLKGYKONEAIFSPLEBDVPEDAQEM 540
 QY 541 VESGYCEGDHKTAKAIKORVSLIKRRBQROLVREEQEKKQESSLMQOVBQSSASQ 600
 Db 541 VESGYCEGDHKTAKAIKORVSLIKRRBQROLVREEQEKKQESSLMQOVBQSSASQ 600
 QY 601 TGIQOLPSASTGIPASTTSASVSTOVEPEPEADQOQLOQOQPSISIVLSDGTVDSGQ 660
 Db 601 TGIQOLPSASTGIPASTTSASVSTOVEPEPEADQOQLOQOQPSISIVLSDGTVDSGQ 660
 QY 661 SSVTESSVSSQOQVVSQSGOHEOAHSTGTVPGHIPSTVOAQSOHGYVPPSSVAQSGQ 720
 Db 661 SSVTESSVSSQOQVVSQSGOHEOAHSTGTVPGHIPSTVOAQSOHGYVPPSSVAQSGQ 720
 QY 721 QPSSSSILTVASSSQPIQHPQOQOQIQQAPAPQOQVQSLSGQTSSEATTAQVPSQAP 780
 Db 721 QPSSSSILTVASSSQPIQHPQOQOQIQQAPAPQOQVQSLSGQTSSEATTAQVPSQAP 780
 QY 781 QVLPQVSAKOSTQGVSOVAPAEVAVAAQOPATOPTLLASVDSAHSTVASGMSDGENV 840
 Db 781 QVLPQVSAKOSTQGVSOVAPAEVAVAAQOPATOPTLLASVDSAHSTVASGMSDGENV 840
 QY 841 PSSSGRHEGRTTKDHYRKSVSRSRHEKTSRPKRLINVSNGRVRVCEOLETNRKMYT 900
 Db 841 PSSSGRHEGRTTKDHYRKSVSRSRHEKTSRPKRLINVSNGRVRVCEOLETNRKMYT 900
 QY 901 FKFDLDGDNPEELATIMVNDPILAIERESFVDQVREIIEKADMLSEDSVEBEDQGL 960
 Db 901 FKFDLDGDNPEELATIMVNDPILAIERESFVDQVREIIEKADMLSEDSVEBEDQGL 960
 QY 961 ESLOGKDDYGRSSGQKLEGEFKOPIPASSMPOQIGIPSSSLTVYVHSGRRFYVSPVES 1020
 Db 961 ESLOGKDDYGRSSGQKLEGEFKOPIPASSMPOQIGIPSSSLTVYVHSGRRFYVSPVES 1020
 QY 1021 RLRSKVPSEITITVAASTQSPQMLSHSASSLSLQQAASELRRAOMTGPTAPNPF 1080
 Db 1021 RLRSKVPSEITITVAASTQSPQMLSHSASSLSLQQAASELRRAOMTGPTAPNPF 1080
 QY 1081 SHGTPTFPVVPPLSSINGVPTAAATAVPATSSPPNDISTSVYQSEVTVTEEGIAGV 1140
 Db 1081 SHGTPTFPVVPPLSSINGVPTAAATAVPATSSPPNDISTSVYQSEVTVTEEGIAGV 1140
 QY 1141 ATSTGVVTSGLPLTPPVSSEPVLSVSSITIPAVVISTSPISQVNTSSEIVVSTA 1200
 Db 1141 ATSTGVVTSGLPLTPPVSSEPVLSVSSITIPAVVISTSPISQVNTSSEIVVSTA 1200
 QY 1201 LYPSTVATSAAGSGTATGPKPAPVAVSQOAGSTTVGATLNSVSTTTFPSTASDLS 1260
 Db 1201 LYPSTVATSAAGSGTATGPKPAPVAVSQOAGSTTVGATLNSVSTTTFPSTASDLS 1260
 QY 1261 IQLSSSTSTPTLAETTVVVAHSLDKTSHSITGLAFSLAPSSSSSGAGVSSYISQPG 1320

Db 1261 IQLSSSTSTPTLAETTVVVAHSLDKTSHSITGLAFSLAPSSSSSGAGVSSYISQPG 1320
 QY 1321 LHPVLVPSVASTPRLPOAGPSTPLLPVPSIPPLVQVAVANVPAVQOQLIHQOPAL 1380
 Db 1321 LHPVLVPSVASTPRLPOAGPSTPLLPVPSIPPLVQVAVANVPAVQOQLIHQOPAL 1380
 QY 1381 LHPQPTHCEVSDTOPKAPGIDDIKLEBKRLSIFSEHSSGAGASVLESTLVIES 1440
 Db 1381 LHPQPTHCEVSDTOPKAPGIDDIKLEBKRLSIFSEHSSGAGASVLESTLVIES 1440
 QY 1441 TTPPGIPTTAAVAPSKLITSTSTCLPPTNIPLGVALPVTPVVTQGVSTIPVSTTSYK 1500
 Db 1441 TTPPGIPTTAAVAPSKLITSTSTCLPPTNIPLGVALPVTPVVTQGVSTIPVSTTSYK 1500
 QY 1501 PGTAPSKPPLTKAPVLVPGVELPAGTLPSEQLPFPFPSPSTQSQOQPLEDDAQLRRLTSP 1560
 Db 1501 PGTAPSKPPLTKAPVLVPGVELPAGTLPSEQLPFPFPSPSTQSQOQPLEDDAQLRRLTSP 1560
 QY 1561 EXITTVSANGPVSMAPALTEAGTOPKGVSVQKGPVLATSSGAGVFKMGFPVSA 1620
 Db 1561 EXITTVSANGPVSMAPALTEAGTOPKGVSVQKGPVLATSSGAGVFKMGFPVSA 1620
 QY 1621 DGAQKEGKNSDEDAKSVHFESSSTSESVLSSSPSESTLVKPEPVGITIPGISDVPSAH 1680
 Db 1621 DGAQKEGKNSDEDAKSVHFESSSTSESVLSSSPSESTLVKPEPVGITIPGISDVPSAH 1680
 QY 1681 KTTASEAKSDTGOPTKGRFOVTTTANKVGRFVSKEDEKITDTKKEGPVAPPFMDLEQ 1740
 Db 1681 KTTASEAKSDTGOPTKGRFOVTTTANKVGRFVSKEDEKITDTKKEGPVAPPFMDLEQ 1740
 QY 1741 AVLPVAVIPKKEBELSPSHLNGPSSDPEAFLSRVDGSGSPHSQSLSPQNTL 1800
 Db 1741 AVLPVAVIPKKEBELSPSHLNGPSSDPEAFLSRVDGSGSPHSQSLSPQNTL 1800
 QY 1801 SOSLSNSFNSSNSWSDNESDIEDDELRLRLRDKLKEIODLSQSKHIESELTYKLG 1860
 Db 1801 SOSLSNSFNSSNSWSDNESDIEDDELRLRLRDKLKEIODLSQSKHIESELTYKLG 1860
 QY 1861 KVPRAVILPPAAPLJSGRRRPTKSGKSSRSSLSLGNKSPQLSGQSAASVLPFOQT 1920
 Db 1861 KVPRAVILPPAAPLJSGRRRPTKSGKSSRSSLSLGNKSPQLSGQSAASVLPFOQT 1920
 QY 1921 LHPGNIPEBGOQLQPLKPSBSSDNLGYAFTSDGAI SVPSLSAPQGIKOPSSSXN 1978
 Db 1921 LHPGNIPEBGOQLQPLKPSBSSDNLGYAFTSDGAI SVPSLSAPQGIKOPSSSXN 1978

RESULT 14
 US-09-854-856-10
 : Sequence 10, Application US/09854856
 : Patent No. 6541252
 : GENERAL INFORMATION:
 : APPLICANT: Walke, D. Wade
 : APPLICANT: Halbun, Erin
 : APPLICANT: Donoho, Gregory
 : APPLICANT: Turner, C. Alexander Jr.
 : TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
 : TITLE OF INVENTION: Encoding the Same
 : FILE REFERENCE: LEX-0178-USA
 : CURRENT APPLICATION NUMBER: US/09/854,856
 : PRIOR FILING DATE: 2001-05-14
 : PRIOR FILING DATE: 2000-05-19
 : NUMBER OF SEQ ID NOS: 64
 : SOFTWARE: PasteSeq for Windows Version 4.0
 : SEQ ID NO 10
 : LENGTH: 2092
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: VARIANT
 : LOCATION: (1)...(2092)

OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-10

Query Match 91.3%; Score 9872.5; DB 4; Length 2092;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1972; Conservative 1; Mismatches 1; Indels 97; Gaps 2;

QY 1 MSGAAEKOSSPTGSLFLSPAPAPKNGSSSDSSVGEKLGAAADATVATGRTEERRRHT 60
DB 1 MSGAAEKOSSPTGSLFLSPAPAPKNGSSSDSSVGEKLGAAADATVATGRTEERRRHT 60
QY 61 MDKDSRGAATTTTTEHFRFRSVICDSNATALELPGHPLSLPQPSIPAAPQSAPEPH 120
DB 61 MDKDSRGAATTTTTEHFRFRSVICDSNATALELPGHPLSLPQPSIPAAPQSAPEPH 120
QY 121 REEVTATATSOVAAQPPAAAAPGQAVAGAPSTVPSSTKDRPVSOPLSVSKEEPP 180
DB 121 REEVTATATSOVAAQPPAAAAPGQAVAGAPSTVPSSTKDRPVSOPLSVSKEEPP 180
QY 181 ARSGGGGSAKEPQBERSQODDIELEETKAVGMSNDRFLKFDIEIGRGSFKTVYKGLD 240
DB 181 ARSGGGGSAKEPQBERSQODDIELEETKAVGMSNDRFLKFDIEIGRGSFKTVYKGLD 240
QY 241 TETTVAVAMCELOQRKLTKEERQRFKEBAEMLKGLQHPNIVRFYDSWESTVKGKCTIV 300
DB 241 TETTVAVAMCELOQRKLTKEERQRFKEBAEMLKGLQHPNIVRFYDSWESTVKGKCTIV 300
QY 301 TELMTSGTLKTYLKRFPVMTKVLRSNCROILKGLQFHTPTPIIHRDLKCDNIFTGP 360
DB 301 TELMTSGTLKTYLKRFPVMTKVLRSNCROILKGLQFHTPTPIIHRDLKCDNIFTGP 360
QY 361 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPEMYEKEKYDESUVVAFGCMLEMAISEY 420
DB 361 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPEMYEKEKYDESUVVAFGCMLEMAISEY 420
QY 421 PYSECQAAOYRRVTSGVYKPAFDFKVAIPVKEIIGCIRONDERYSIKDLINHAFFQ 480
DB 421 PYSECQAAOYRRVTSGVYKPAFDFKVAIPVKEIIGCIRONDERYSIKDLINHAFFQ 480
QY 481 EETGVRELAEBEDGEEKIATKMLRIDIKLKAKKYKONDAIBSFLEBDVPEVDAQEM 540
DB 481 EETGVRELAEBEDGEEKIATKMLRIDIKLKAKKYKONDAIBSFLEBDVPEVDAQEM 540
QY 541 VESGVYCEGDHKTAKAIKDRVSLIKRREBOROLVREOEKKKESSSLKQOVVQSSASQ 600
DB 541 VESGVYCEGDHKTAKAIKDRVSLIKRREBOROLVREOEKKKESSSLKQOVVQSSASQ 600
QY 601 TGIKOLPSASTGIPASTTSASVSTQVEPEPEADQHQLOQYQOQPSISVLSDTGTVDSGQ 660
DB 601 TGIKOLPSASTGIPASTTSASVSTQVEPEPEADQHQLOQYQOQPSISVLSDTGTVDSGQ 660
QY 661 SSVFTESRVSSQOVYSGSQHEQAHSTGYVGHIPSTVQAOSQPHGYVPPSSVAQGSQ 720
DB 661 SSVFTESRVSSQOVYSGSQHEQAHSTGYVGHIPSTVQAOSQPHGYVPPSSVAQGSQ 720
QY 721 QPSSSLTGVSSSQPIQHPQOQOGIIOOTAPPOQTVQYSLSTQTSSEATTAAPQSPQAP 780
DB 721 QPSSSLTGVSSSQPIQHPQOQOGIIOOTAPPOQTVQYSLSTQTSSEATTAAPQSPQAP 780
QY 781 QVLPVVSAGQ-----791
DB 781 QVLPVVSAGQ-----791
QY 792 -----STQGVQVAPAEVAV 807
DB 792 -----STQGVQVAPAEVAV 807
QY 841 VVQPYVESNLLVPMGVGVQVQVQPGGSLAQAPITSSQQAIVLESTQGVQVAPAEVAV 900
DB 841 VVQPYVESNLLVPMGVGVQVQVQPGGSLAQAPITSSQQAIVLESTQGVQVAPAEVAV 900
QY 808 ACPQATQPTTLASVDSAHSDVASGMSDGNENYSSSGRHEGRTTKRHYKSVRSRHE 867
DB 808 ACPQATQPTTLASVDSAHSDVASGMSDGNENYSSSGRHEGRTTKRHYKSVRSRHE 867
QY 901 ACPQATQPTTLASVDSAHSDVASGMSDGNENYSSSGRHEGRTTKRHYKSVRSRHE 960
DB 901 ACPQATQPTTLASVDSAHSDVASGMSDGNENYSSSGRHEGRTTKRHYKSVRSRHE 960
QY 868 KTSRPKRLINVSNKGDVVEQLETHNRMTVFKEDLDGDNPEEIAITIMVNDPILALE 927
DB 868 KTSRPKRLINVSNKGDVVEQLETHNRMTVFKEDLDGDNPEEIAITIMVNDPILALE 927

DB 961 KTSRPKRLINVSNKGDVVEQLETHNRMTVFKEDLDGDNPEEIAITIMVNDPILALE 1020
QY 928 RESFVDQREILIEKADENLSEDAVEPEQDGLSLQCKDDYGFSGSKLEGEFQPIPA 987
DB 1021 RESFVDQREILIEKADENLSEDAVEPEQDGLSLQCKDDYGFSGSKLEGEFQPIPA 1080
QY 988 SSMFQOIGIPSSLTQVYHSGRRFIVSPVESRLRESKVPBSITDTVAASTQSPGMN 1047
DB 1081 SSMFQOIGIPSSLTQVYHSGRRFIVSPVESRLRESKVPBSITDTVAASTQSPGMN 1140
QY 1048 LSHGASSLSQQAASELRRAQMTGPTAPNPGHTGTFPVVPEPFLSSIAIGVPTAAT 1107
DB 1141 LSHGASSLSQQAASELRRAQMTGPTAPNPGHTGTFPVVPEPFLSSIAIGVPTAAT 1200
QY 1108 APVPATSSPNDISTSVIYQSEVTVPEEGIAGVATSTGVVTSQGLPIPPVSESPVLSVV 1167
DB 1201 APVPATSSPNDISTSVIYQSEVTVPEEGIAGVATSTGVVTSQGLPIPPVSESPVLSVV 1260
QY 1168 SSITIPAVSISTSPLOVPTSTSEIVVSTALYPSVTVSATASAGGSTATGPRPA 1227
DB 1261 SSITIPAVSISTSPLOVPTSTSEIVVSTALYPSVTVSATASAGGSTATGPRPA 1320
QY 1228 VVSQQAAGSTTVGATLTSVSTTSFPSTASQSLSTQSSSTPTLAETVVVSAHSLDKTS 1287
DB 1321 VVSQQAAGSTTVGATLTSVSTTSFPSTASQSLSTQSSSTPTLAETVVVSAHSLDKTS 1380
QY 1288 HSSITGLAFSLAPSSSSSPGAGVSYISQPGIHLPLVPSVIASTPILPQAGPSTPPL 1347
DB 1381 HSSITGLAFSLAPSSSSSPGAGVSYISQPGIHLPLVPSVIASTPILPQAGPSTPPL 1440
QY 1348 LPQVPSIPPLVOPANPAVQOQLIHSQOPALLPQPHHCPEVDSDTQPKAGIDIDIX 1407
DB 1441 LPQVPSIPPLVOPANPAVQOQLIHSQOPALLPQPHHCPEVDSDTQPKAGIDIDIX 1500
QY 1408 TLEBKRLSEHSSSSGAQASVLETSVLESTVTEGIPPTAASPRLTSTSTCLPP 1467
DB 1501 TLEBKRLSEHSSSSGAQASVLETSVLESTVTEGIPPTAASPRLTSTSTCLPP 1560
QY 1468 TNLPLGTVALPVPVTPGVSTPVTSTTSGVKRGAPSKRPPLTKADVLVGTLEPAGTL 1527
DB 1561 TNLPLGTVALPVPVTPGVSTPVTSTTSGVKRGAPSKRPPLTKADVLVGTLEPAGTL 1620
QY 1528 PSEQLPFPPEPSSLTQSQOPLIEDLDQRLTSLPEXITVTSVANGPVSAAPALTEACTOP 1587
DB 1621 PSEQLPFPPEPSSLTQSQOPLIEDLDQRLTSLPEXITVTSVANGPVSAAPALTEACTOP 1680
QY 1588 QKGVSYQVEGPVLATSGAGVFKRGFPVVAADGAQKEGKNKSEDAKSVHPESTSESS 1647
DB 1681 QKGVSYQVEGPVLATSGAGVFKRGFPVVAADGAQKEGKNKSEDAKSVHPESTSESS 1740
QY 1648 VLSSSPEESTLVKPEPNGITIPGISDVPBSAHKTTASEAKSDTGQTKVGRFOVTTAN 1707
DB 1741 VLSSSPEESTLVKPEPNGITIPGISDVPBSAHKTTASEAKSDTGQTKVGRFOVTTAN 1800
QY 1708 KYGRFYSVKTEKIDITTKKGPVAPSPFMDLEAVIPAVIPKKEKPELSPSHLNGSSD 1767
DB 1801 KYGRFYSVKTEKIDITTKKGPVAPSPFMDLEAVIPAVIPKKEKPELSPSHLNGSSD 1860
QY 1768 PEAAPLSRDVDDSGSPHSPHOLSKSLPSQNLQSLSNSFNSSYMSSDNESEDIEDLX 1827
DB 1861 PEAAPLSRDVDDSGSPHSPHOLSKSLPSQNLQSLSNSFNSSYMSSDNESEDIEDLX 1920
QY 1828 LELRLRLDKHLKAIQLOLSQKHEISLTKLGVKPPAVIIPPAAPLSGRRRPTYSKGS 1887
DB 1921 LELRLRLDKHLKAIQLOLSQKHEISLTKLGVKPPAVIIPPAAPLSGRRRPTYSKGS 1980
QY 1888 KSSRSSSLGKNSQOLSGNLSGQSAASVILHQOQLLHPGNIPESGQOQLLOPLKPSSSDN 1947
DB 1981 KSSRSSSLGKNSQOLSGNLSGQSAASVILHQOQLLHPGNIPESGQOQLLOPLKPSSSDN 2040
QY 1948 LYSAPTSQALISVPSLSAPGQG-----TSSSTN 1974
DB 2041 LYSAPTSQALISVPSLSAPGQG-----TSSSTN 2071

RESULT 15
US-09-854-856-4
Sequence 4, Application US/09854856
Patent No. 6541252
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Hilbun, Erin
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
FILE REFERENCE: LEX-0178-USA
CURRENT APPLICATION NUMBER: US/09/854,856
CURRENT FILING DATE: 2001-05-14
PRIORITY APPLICATION NUMBER: US 60/206,015
PRIORITY FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 2245
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)-(2245)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-4

Query Match 90.6%; Score 9796; DB 4; Length 2245;
Best Local Similarity 88.7%; Pred. No. 0;
Matches 1972; Conservative 1; Mismatches 1; Indels 250; Gaps 2;

QY 1 MSGGAAGKOSTPGSLSPAPAPKNGSSSDSVGKLGAAADAYTGTETRRRRHT 60
DB 1 MSGGAAGKOSTPGSLSPAPAPKNGSSSDSVGKLGAAADAYTGTETRRRRHT 60
QY 61 MDKSRBAATTTTTERFPFRSVICDSNATLELPGPLSLPQPSIPAAVPOGAPPEPH 120
DB 61 MDKSRBAATTTTTERFPFRSVICDSNATLELPGPLSLPQPSIPAAVPOGAPPEPH 120
QY 121 REETVATATISQVAAQPPAAAPGEOAVAGPAPSTVSTSKDRPVSOPLVSGKEBPP 180
DB 121 REETVATATISQVAAQPPAAAPGEOAVAGPAPSTVSTSKDRPVSOPLVSGKEBPP 180
QY 181 ARSGSGGSAKEPOEBERSQODDIEBELTKAVGMSNDGRFLKPIETIGRSGFKTVYKGLD 240
DB 181 ARSGSGGSAKEPOEBERSQODDIEBELTKAVGMSNDGRFLKPIETIGRSGFKTVYKGLD 240
QY 241 TETTVAVWMCLOBRKLTSEKORFKERAEMLKGLPHNIVRFYDSWSTVKGKCIYLV 300
DB 241 TETTVAVWMCLOBRKLTSEKORFKERAEMLKGLPHNIVRFYDSWSTVKGKCIYLV 300
QY 301 TELMTSGTLTKYLRFKYMKIKVLRSWCQILKGLQFLHTRTPPIIHRDLKCNIFITGP 360
DB 301 TELMTSGTLTKYLRFKYMKIKVLRSWCQILKGLQFLHTRTPPIIHRDLKCNIFITGP 360
QY 361 TGSVKTIGDLATLKRPASFAKSVIGTPFPMAPEMTEEEKYDSVDYVAGCMLEMAISEY 420
DB 361 TGSVKTIGDLATLKRPASFAKSVIGTPFPMAPEMTEEEKYDSVDYVAGCMLEMAISEY 420
QY 421 PYSECQNAOQYRRVTSVGPASFPKVAIPVEKTEIEGCIKONKDERYSIDOLNHAFFQ 480
DB 421 PYSECQNAOQYRRVTSVGPASFPKVAIPVEKTEIEGCIKONKDERYSIDOLNHAFFQ 480
QY 481 EETGVRVLAEDDEGEKIAIKLWLRIEDIKKLGKXYKDNEMAIERSFDLERDVPEDVAGEM 540
DB 481 EETGVRVLAEDDEGEKIAIKLWLRIEDIKKLGKXYKDNEMAIERSFDLERDVPEDVAGEM 540
QY 541 VESGVGCGDHKTKAKIKRVSILIKRKGEOQLVREBQEKKKQVESSLKQOVESGASQ 600
DB 541 VESGVGCGDHKTKAKIKRVSILIKRKGEOQLVREBQEKKKQVESSLKQOVESGASQ 600

QY 601 TGIKOLPSASTGIPNASTTSASVSTQVEPEEPADHQOLOIQOQPSISVLSDETVDGSG 660
DB 601 TGIKOLPSASTGIPNASTTSASVSTQVEPEEPADHQOLOIQOQPSISVLSDETVDGSG 660
QY 661 SSVFTSRVSSQCTVYSGQHEQAHSTGVPHGIPSTVQAQSPHGVYPPSSVAQSQSG 720
DB 661 SSVFTSRVSSQCTVYSGQHEQAHSTGVPHGIPSTVQAQSPHGVYPPSSVAQSQSG 720
QY 721 QPSSSLTGVSASSQPIQHPOQOQIGIQTAPPOQTVYSLQTSSEATTAQVSOQAP 780
DB 721 QPSSSLTGVSASSQPIQHPOQOQIGIQTAPPOQTVYSLQTSSEATTAQVSOQAP 780
QY 781 QVLPQVSAGKO----- 791
DB 781 QVLPQVSAGKO----- 791
QY 792 ----- 791
DB 792 ----- 791
QY 841 VSGIPSTPHVSTAQTGFSPLITMAAGITOPLLTASATTAIEGVSTVPSQLEPTL 900
DB 841 VSGIPSTPHVSTAQTGFSPLITMAAGITOPLLTASATTAIEGVSTVPSQLEPTL 900
QY 901 QPVTQLPSQVHPQLQPAVQSMGIPANLGOAAEVLPSGQVLYXGPPRRLPQYPGNSNI 960
DB 901 QPVTQLPSQVHPQLQPAVQSMGIPANLGOAAEVLPSGQVLYXGPPRRLPQYPGNSNI 960
QY 992 ----- 791
DB 992 ----- 791
QY 961 APSSNVAVCIHSTVLXPMPTREVLATPGYFTVQPVYESNLVPMGVGQVQVSPG 1020
DB 961 APSSNVAVCIHSTVLXPMPTREVLATPGYFTVQPVYESNLVPMGVGQVQVSPG 1020
QY 792 -----STOGVQVAPAEVVAQAQOATOPTTLASVDSASHSVASGMS 834
DB 792 -----STOGVQVAPAEVVAQAQOATOPTTLASVDSASHSVASGMS 834
QY 1021 GSIAQAPTTSSQQAVALSTQGVQVAPAEVVAQAQOATOPTTLASVDSASHSVASGMS 1080
DB 1021 GSIAQAPTTSSQQAVALSTQGVQVAPAEVVAQAQOATOPTTLASVDSASHSVASGMS 1080
QY 835 DGENVNPSSGRHEGRITTKHRYKRSVRSRHEKTSRPKRIINVSNKGRVVECOLETH 894
DB 835 DGENVNPSSGRHEGRITTKHRYKRSVRSRHEKTSRPKRIINVSNKGRVVECOLETH 894
QY 1081 DGENVNPSSGRHEGRITTKHRYKRSVRSRHEKTSRPKRIINVSNKGRVVECOLETH 1140
DB 1081 DGENVNPSSGRHEGRITTKHRYKRSVRSRHEKTSRPKRIINVSNKGRVVECOLETH 1140
QY 895 NRKAVTEKFDLDGNDPEIATIMVNDPILAIERESFVQVREIIEKADEMLSEDEVEP 954
DB 895 NRKAVTEKFDLDGNDPEIATIMVNDPILAIERESFVQVREIIEKADEMLSEDEVEP 954
QY 1141 NRKAVTEKFDLDGNDPEIATIMVNDPILAIERESFVQVREIIEKADEMLSEDEVEP 1200
DB 1141 NRKAVTEKFDLDGNDPEIATIMVNDPILAIERESFVQVREIIEKADEMLSEDEVEP 1200
QY 955 EGDQGLSLQKQDYGSGSQKIEGFEKQIPASSMPOQIGIPSSLTQVHSGRPFIV 1014
DB 955 EGDQGLSLQKQDYGSGSQKIEGFEKQIPASSMPOQIGIPSSLTQVHSGRPFIV 1014
QY 1201 EGDQGLSLQKQDYGSGSQKIEGFEKQIPASSMPOQIGIPSSLTQVHSGRPFIV 1260
DB 1201 EGDQGLSLQKQDYGSGSQKIEGFEKQIPASSMPOQIGIPSSLTQVHSGRPFIV 1260
QY 1015 SPVSESRRESKVPSEITDTVAASTAQSPGMMLSHSASSLSLQAAFSERLRAQMTGEPN 1074
DB 1015 SPVSESRRESKVPSEITDTVAASTAQSPGMMLSHSASSLSLQAAFSERLRAQMTGEPN 1074
QY 1261 SPVSESRRESKVPSEITDTVAASTAQSPGMMLSHSASSLSLQAAFSERLRAQMTGEPN 1320
DB 1261 SPVSESRRESKVPSEITDTVAASTAQSPGMMLSHSASSLSLQAAFSERLRAQMTGEPN 1320
QY 1075 TAPNFSHTGTPPEVVPPLSSIAGVPTTAATAAPVATSSPPNDISTSVIQSEVTVPTE 1134
DB 1075 TAPNFSHTGTPPEVVPPLSSIAGVPTTAATAAPVATSSPPNDISTSVIQSEVTVPTE 1134
QY 1321 TAPNFSHTGTPPEVVPPLSSIAGVPTTAATAAPVATSSPPNDISTSVIQSEVTVPTE 1380
DB 1321 TAPNFSHTGTPPEVVPPLSSIAGVPTTAATAAPVATSSPPNDISTSVIQSEVTVPTE 1380
QY 1135 EGIAGVATSTGVNTSGGLPIPVSESPVLSVSVSITIPVAVSITSSPSIQVFTSSEI 1194
DB 1135 EGIAGVATSTGVNTSGGLPIPVSESPVLSVSVSITIPVAVSITSSPSIQVFTSSEI 1194
QY 1381 EGIAGVATSTGVNTSGGLPIPVSESPVLSVSVSITIPVAVSITSSPSIQVFTSSEI 1440
DB 1381 EGIAGVATSTGVNTSGGLPIPVSESPVLSVSVSITIPVAVSITSSPSIQVFTSSEI 1440
QY 1195 VVSSTALPVSVTASATASAGSSTATGPKPRAVVSQQAAGSTTVGATLTVSFTTSPS 1254
DB 1195 VVSSTALPVSVTASATASAGSSTATGPKPRAVVSQQAAGSTTVGATLTVSFTTSPS 1254
QY 1441 VVSSTALPVSVTASATASAGSSTATGPKPRAVVSQQAAGSTTVGATLTVSFTTSPS 1500
DB 1441 VVSSTALPVSVTASATASAGSSTATGPKPRAVVSQQAAGSTTVGATLTVSFTTSPS 1500
QY 1255 TAAQSLISQSSSTPTLAEVTVVVAHSLDKTSHSSTTGAFSLASAPSSSSPAGVSSY 1314
DB 1255 TAAQSLISQSSSTPTLAEVTVVVAHSLDKTSHSSTTGAFSLASAPSSSSPAGVSSY 1314
QY 1501 TAAQSLISQSSSTPTLAEVTVVVAHSLDKTSHSSTTGAFSLASAPSSSSPAGVSSY 1560
DB 1501 TAAQSLISQSSSTPTLAEVTVVVAHSLDKTSHSSTTGAFSLASAPSSSSPAGVSSY 1560
QY 1315 ISQGGHLPVIVSVIASTPILPQAAGPTSPILLPOVPSIPVLYOPVAVVAQOOLTHS 1374
DB 1315 ISQGGHLPVIVSVIASTPILPQAAGPTSPILLPOVPSIPVLYOPVAVVAQOOLTHS 1374
QY 1561 ISQGGHLPVIVSVIASTPILPQAAGPTSPILLPOVPSIPVLYOPVAVVAQOOLTHS 1620
DB 1561 ISQGGHLPVIVSVIASTPILPQAAGPTSPILLPOVPSIPVLYOPVAVVAQOOLTHS 1620
QY 1375 QPOPALLPNOCHTCEVSDTOPKACGIDIDIKLLEKRLSLFSEHSSSGAQAHSVSLT 1434
DB 1375 QPOPALLPNOCHTCEVSDTOPKACGIDIDIKLLEKRLSLFSEHSSSGAQAHSVSLT 1434
QY 1621 QPOPALLPNOCHTCEVSDTOPKACGIDIDIKLLEKRLSLFSEHSSSGAQAHSVSLT 1680
DB 1621 QPOPALLPNOCHTCEVSDTOPKACGIDIDIKLLEKRLSLFSEHSSSGAQAHSVSLT 1680

QY	1435	SLVIESITVPGIPTTAVAPSKLITSTSTCUPPTNLPLGTVALPVTPTVTPQGVSTPVST	1494
Db	1681	SLVIESITVPGIPTTAVAPSKLITSTSTCUPPTNLPLGTVALPVTPTVTPQGVSTPVST	1740
QY	1495	TTSGVYKPGTAPSKPPLITAPVLPVGTBLPAGTLPEEQULPPPPGSLTQSQOPLIEDLDAQL	1554
Db	1741	TTSGVYKPGTAPSKPPLITAPVLPVGTBLPAGTLPEEQULPPPPGSLTQSQOPLIEDLDAQL	1800
QY	1555	RRTLSPEKXITVTSAGVPVMAAPTAITEAGTQPOKGVQVQVEGVPYATSSGAGVFKMGFR	1614
Db	1801	RRTLSPEKXITVTSAGVPVMAAPTAITEAGTQPOKGVQVQVEGVPYATSSGAGVFKMGFR	1860
QY	1615	QVSVAADGAQKEGKNKSEDAKSVHESSTSSSVLSSSSPESTLVKPPENGITTPGISSD	1674
Db	1861	QVSVAADGAQKEGKNKSEDAKSVHESSTSSSVLSSSSPESTLVKPPENGITTPGISSD	1920
QY	1675	VPESAHTTASEAKSDTQPTKVGRFQVTTTANKVGRFVSCTEDKITDTYKKEGVPVSP	1734
Db	1921	VPESAHTTASEAKSDTQPTKVGRFQVTTTANKVGRFVSCTEDKITDTYKKEGVPVSP	1980
QY	1735	FMDLEQAVLPVAVIPKKEPELSEPSHNGPSSDPEAFLSRDVPDGGSGPHSPHOLSKS	1794
Db	1981	FMDLEQAVLPVAVIPKKEPELSEPSHNGPSSDPEAFLSRDVPDGGSGPHSPHOLSKS	2040
QY	1795	LPSQNLQSLSNSFNSSSYMSSDNESDIEDDLKLELRRLRDKHKEIQDLQSRQKHIES	1854
Db	2041	LPSQNLQSLSNSFNSSSYMSSDNESDIEDDLKLELRRLRDKHKEIQDLQSRQKHIES	2100
QY	1855	LYTKLGKVPVPPVILPPAAPLSGRRRPTKSGSKSSRSSSLGNKSPOLSGNLSCQSAASV	1914
Db	2101	LYTKLGKVPVPPVILPPAAPLSGRRRPTKSGSKSSRSSSLGNKSPOLSGNLSCQSAASV	2160
QY	1915	LHPQOTLHPGNIPEGONQLOPLKPSPSGDNLYSAFTSDGATSVPSLSAPQGG---T	1970
Db	2161	LHPQOTLHPGNIPEGONQLOPLKPSPSGDNLYSAFTSDGATSVPSLSAPQGG---T	2220
QY	1971	SSTN 1974	
Db	2221	SSKN 2224	

Search completed: September 24, 2004, 01:12:44
 Job time : 64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2004, 00:51:58 ; Search time 189 Seconds

(without alignments)
3193.235 Million cell updates/sec

Title: US-10-010-720-14

Perfect score: 10812
Sequence: 1 MSGGAARQSTPGSLFLSP.....NISMLQKISNPSPGSLNRTT 2136

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_29Jan04:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2002s:*
7: geneseqp2003as:*
8: geneseqp2003bs:*
9: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10810	100.0	2136	7	ABO44387 Novel hum
2	10794.5	99.8	2135	5	AAE21714 Human PKI
3	10753.5	99.5	2229	7	ABO44384 Novel hum
4	10677	98.8	2382	6	ABP71619 Human MNK
5	10677	98.8	2382	7	ABO44381 Novel hum
6	10646	98.5	2108	7	ABO44395 Novel hum
7	10646	98.5	2108	7	ADCG9049 Novel hum
8	10589.5	97.9	2201	7	ABO44392 Novel hum
9	10513	97.2	2354	7	ABO44389 Novel hum
10	10508	97.2	2076	7	ABO44403 Novel hum
11	10451.5	96.7	2169	7	ABO44400 Novel hum
12	10375	96.0	2332	7	ABO44397 Novel hum
13	10344	95.7	2048	7	ABO44411 Novel hum
14	10287.5	95.1	2141	7	ABO44408 Novel hum
15	10211	94.4	2294	7	ABO44405 Novel hum
16	9929	91.8	1999	7	ABO44388 Novel hum
17	9872.5	91.3	2092	7	ABO44385 Novel hum
18	9766	90.6	2245	7	ABO44382 Novel hum
19	9765	90.3	1971	7	ABO44386 Novel hum
20	9708.5	89.8	2064	7	ABO44393 Novel hum
21	9632	89.1	2217	7	ABO44390 Novel hum
22	9627	89.0	1939	7	ABO44404 Novel hum
23	9612.5	88.9	1920	4	AA85656 Novel pro
24	9570.5	88.5	2032	7	ABO44401 Novel hum
25	9434	87.8	2185	7	ABO44398 Novel hum

26	9463	87.5	1911	7	ABO44412
27	9406.5	87.0	2004	7	ABO44409
28	9330	86.3	2157	7	ABO44406
29	4053	37.5	982	7	ABO44383
30	4043	37.4	829	7	ABO44386
31	3889	36.0	954	7	ABO44391
32	3879	35.9	801	7	ABO44394
33	3751	34.7	922	7	ABO44399
34	3741	34.6	769	7	ABO44402
35	3587	33.2	894	7	ABO44407
36	3577	33.1	741	7	ABO44410
37	2903	26.8	565	4	AA85785
38	2631	24.3	2193	6	ABR4219
39	2524	23.3	502	4	AA838736
40	2469.5	22.8	591	4	AA838737
41	2419	22.4	470	4	AA840522
42	2419	22.4	470	4	AA840523
43	2130.5	19.7	1345	5	AAE25097
44	2126	19.7	1243	6	ABP71620
45	2126	19.7	1513	4	AAU03531

ALIGNMENTS

ABO44412 Novel hum
ABO44409 Novel hum
ABO44406 Novel hum
ABO44383 Novel hum
ABO44386 Novel hum
ABO44391 Novel hum
ABO44394 Novel hum
ABO44399 Novel hum
ABO44402 Novel hum
ABO44407 Novel hum
ABO44410 Novel hum
ABO44410 Novel hum
ABR4219 Human kin
AA838736 Human pol
AA838737 Human pol
AA840522 Human pol
AA840523 Human pol
AAE25097 Human kin
ABP71620 Human MNK
AAU03531 Human pro

RESULT 1
ID ABO44387 standard; protein; 2136 AA.
AC ABO44387;
DT 26-SEP-2003 (first entry)
DE Novel human protein kinase #7.
KW Human; kinase; enzyme; cosmetic application; nutraceutical application.
OS Homo sapiens.
PN US6541252-B1.
PD 01-APR-2003.
PE 14-MAY-2001; 2001US-00854856.
PF 19-MAY-2000; 2000US-0206015P.
PR (LEXI-) LEXICON GENETICS INC.
PS Walke DW, Hilbun E, Donoho G, Turner CA;
DR WPI; 2003-575927/54.
DR N-PSDB; ACH03770.
XX New nucleic acid encoding novel human proteins, useful in cosmetic and nutraceutical applications.
PT nutraceutical applications.
PS Claim 1, Page; 11pp; English.
CC The invention relates to a new isolated nucleic acid encoding a novel human protein kinase. The nucleic acid is useful in cosmetic and nutraceutical applications. The present sequence represents the amino acid sequence of a novel human protein kinase. Note: The sequence data for this patent did not from part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=6541252B1

SO Sequence 2136 AA;
Query Match 100.0%; Score 10810; DB 7; Length 2136;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGGAABKOSSTGSLPLSPAPAPKNGSSSDSSVGEKLGAAADAVTGTTEERYRRRT 60
 Db 1 MSGGAABKOSSTGSLPLSPAPAPKNGSSSDSSVGEKLGAAADAVTGTTEERYRRRT 60
 QY 61 MDKOSGAATTTTTERFRFRFRSVI CDSNATALEPLPLPLPOPSPAPVPOSAPPEPH 120
 Db 61 MDKOSGAATTTTTERFRFRFRSVI CDSNATALEPLPLPLPOPSPAPVPOSAPPEPH 120
 QY 121 REETVATATSVQAQOPPAAPAEQAVAGPADSVTSSTSKDRPVSOPLVSKKEPP 180
 Db 121 REETVATATSVQAQOPPAAPAEQAVAGPADSVTSSTSKDRPVSOPLVSKKEPP 180
 QY 181 ARSGSGGSKKEPOEERSQOQDDIELETKAVGMSNGRFLKFDIEIGRSFPTVYGLD 240
 Db 181 ARSGSGGSKKEPOEERSQOQDDIELETKAVGMSNGRFLKFDIEIGRSFPTVYGLD 240
 QY 241 TETVEVAMELODRKLTKEERQFKEEAEMLKGLQHPNIVRFYDSESTVKGKCIVLV 300
 Db 241 TETVEVAMELODRKLTKEERQFKEEAEMLKGLQHPNIVRFYDSESTVKGKCIVLV 300
 QY 301 TELMTSGTLKTYLKRFRVOMKI KVLRSWCROILKGLQFLHTRTPPI IHRDLKCNIFITGP 360
 Db 301 TELMTSGTLKTYLKRFRVOMKI KVLRSWCROILKGLQFLHTRTPPI IHRDLKCNIFITGP 360
 QY 361 TGSVKIGDLGLATLKRAFPKSVI GTEPEFAPAPEMEYEEKYDESVDVYAFGCMLEMATSEY 420
 Db 361 TGSVKIGDLGLATLKRAFPKSVI GTEPEFAPAPEMEYEEKYDESVDVYAFGCMLEMATSEY 420
 QY 421 PYSBCQNAOQIYRRVTSVGVPAFDPKVAIPEVKEIIBGCIRONKDERYSIKDLNHAFFQ 480
 Db 421 PYSBCQNAOQIYRRVTSVGVPAFDPKVAIPEVKEIIBGCIRONKDERYSIKDLNHAFFQ 480
 QY 481 BETGVAVELAEEDDEKCIATKLMRIDIKLKKYKONENAIERSFLBERDVEDVAQEM 540
 Db 481 BETGVAVELAEEDDEKCIATKLMRIDIKLKKYKONENAIERSFLBERDVEDVAQEM 540
 QY 541 VESGYVEGHHKMAKAIKORVSLIKRREOROLVREOEKKOEESLKOQVOSASQ 600
 Db 541 VESGYVEGHHKMAKAIKORVSLIKRREOROLVREOEKKOEESLKOQVOSASQ 600
 QY 601 TGIKOLPSASTGIPITASVTSASVTOVEPEPEADQOOLQYOQPSISVLDGTVDGQG 660
 Db 601 TGIKOLPSASTGIPITASVTSASVTOVEPEPEADQOOLQYOQPSISVLDGTVDGQG 660
 QY 661 SSVFTESRVSQOQVTSYGSOQHECAHSTGTVPGHIPSTVOAQSOQPHGVYPPSSVAQSOQ 720
 Db 661 SSVFTESRVSQOQVTSYGSOQHECAHSTGTVPGHIPSTVOAQSOQPHGVYPPSSVAQSOQ 720
 QY 721 QPSSSSLTGVSSQPIQHPOQOQSIQQTAPPOQVQYSLSTQSTSSSEITTAQPIVSQCAP 780
 Db 721 QPSSSSLTGVSSQPIQHPOQOQSIQQTAPPOQVQYSLSTQSTSSSEITTAQPIVSQCAP 780
 QY 781 QVLPOVSAKOSTQGVSAPEVAVAVAOQATQPTTLASSVDASHSDVAGSGDGENV 840
 Db 781 QVLPOVSAKOSTQGVSAPEVAVAVAOQATQPTTLASSVDASHSDVAGSGDGENV 840
 QY 841 PSSSGRHRGRTTKKHRYKSVASRSRHEKTSRPKRLINVSNGRVRVCEQJETHNRKVT 900
 Db 841 PSSSGRHRGRTTKKHRYKSVASRSRHEKTSRPKRLINVSNGRVRVCEQJETHNRKVT 900
 QY 901 FKPLDGNPEEIIATIMVNDPFIILAIERESFVDQREIIEKADEMLSEDEVVEEGDGL 960
 Db 901 FKPLDGNPEEIIATIMVNDPFIILAIERESFVDQREIIEKADEMLSEDEVVEEGDGL 960
 QY 961 ESLOGKDYFSGSQKLEGEFKOPIPASSMPQOIGIPITSLTOVHSAGRFIVSPVES 1020
 Db 961 ESLOGKDYFSGSQKLEGEFKOPIPASSMPQOIGIPITSLTOVHSAGRFIVSPVES 1020
 QY 1021 RLRSKVPFSSITTTVAASTAQSPGAMTSHSASSISLQOAFSELRAQMTGPNAPNF 1080
 Db 1021 RLRSKVPFSSITTTVAASTAQSPGAMTSHSASSISLQOAFSELRAQMTGPNAPNF 1080
 QY 1081 SHGPTFPVVPPLSSINGVPTTAAATAPVPATSSPNDISTSVIQSEVTVTEEGIAGV 1140

Db 1081 SHGPTFPVVPPLSSINGVPTTAAATAPVPATSSPNDISTSVIQSEVTVTEEGIAGV 1140
 QY 1141 ANSTGVVTSQGLPIPVSESPVLSVYSSITTPAVVISITSSPSLQVPTSTSEIYVSTA 1200
 Db 1141 ANSTGVVTSQGLPIPVSESPVLSVYSSITTPAVVISITSSPSLQVPTSTSEIYVSTA 1200
 QY 1201 LYPSTVVSATASAGGSTATPGKPPAVVQQAAGSTTVVATLTSVTTSPSTASQLS 1260
 Db 1201 LYPSTVVSATASAGGSTATPGKPPAVVQQAAGSTTVVATLTSVTTSPSTASQLS 1260
 QY 1261 IOLSSSTPTTAEIYVVSASHSDKTSHTSTGLAFSLAPSSSSSPGAGVSSYISQPG 1320
 Db 1261 IOLSSSTPTTAEIYVVSASHSDKTSHTSTGLAFSLAPSSSSSPGAGVSSYISQPG 1320
 QY 1321 LHPVLVPSVASTPPIIPOAAGPSTPLPOVBSIPPLVQVAVNAVPOQTLHSQOPAL 1380
 Db 1321 LHPVLVPSVASTPPIIPOAAGPSTPLPOVBSIPPLVQVAVNAVPOQTLHSQOPAL 1380
 QY 1381 LHPVLPVSVASTPPIIPOAAGPSTPLPOVBSIPPLVQVAVNAVPOQTLHSQOPAL 1440
 Db 1381 LHPVLPVSVASTPPIIPOAAGPSTPLPOVBSIPPLVQVAVNAVPOQTLHSQOPAL 1440
 QY 1441 TVTGPJPTTAVAPSKLITSTTCLPPTNLPICTVALPVTVPVTPQVSTPVSTTSVGV 1500
 Db 1441 TVTGPJPTTAVAPSKLITSTTCLPPTNLPICTVALPVTVPVTPQVSTPVSTTSVGV 1500
 QY 1501 PGTAPEPPLTKAPVLPVGTIELPAGTLPSPOLPPFPBPSTLQSOQPLEDDAOLRRLLSP 1560
 Db 1501 PGTAPEPPLTKAPVLPVGTIELPAGTLPSPOLPPFPBPSTLQSOQPLEDDAOLRRLLSP 1560
 QY 1561 EXITVTSVAVPVSMAAPTAITEAGTOPKGVSOQVEGPVLATSGAGVFMKGFQVSA 1620
 Db 1561 EXITVTSVAVPVSMAAPTAITEAGTOPKGVSOQVEGPVLATSGAGVFMKGFQVSA 1620
 QY 1621 DGAQKEGKNSEDAKSVHFESSSTSESSVSSSPBSTLYKPEPNGITIPGISVPEPSAH 1680
 Db 1621 DGAQKEGKNSEDAKSVHFESSSTSESSVSSSPBSTLYKPEPNGITIPGISVPEPSAH 1680
 QY 1681 KTTASEKASTQGPRTKGRFOVTTTANKVRFVSUKTEKITTPKKGPAVSPFMDLQ 1740
 Db 1681 KTTASEKASTQGPRTKGRFOVTTTANKVRFVSUKTEKITTPKKGPAVSPFMDLQ 1740
 QY 1741 AVLPVAVIPKKEKPELSPSHLNGSSDPEAFLSRVDVDSGSPHSFHOSSKSLPSONL 1800
 Db 1741 AVLPVAVIPKKEKPELSPSHLNGSSDPEAFLSRVDVDSGSPHSFHOSSKSLPSONL 1800
 QY 1801 SOSLSNSFNSSWYSSDNESEDIEDDLKLELRRLADKHLKEIODIQRKHEIBSLYTKLG 1860
 Db 1801 SOSLSNSFNSSWYSSDNESEDIEDDLKLELRRLADKHLKEIODIQRKHEIBSLYTKLG 1860
 QY 1861 KVPRAVILPPAALPSGRRRRTTKSGSKSRSSSLGKNSPOLSGNLSGQSAASVILHPOQT 1920
 Db 1861 KVPRAVILPPAALPSGRRRRTTKSGSKSRSSSLGKNSPOLSGNLSGQSAASVILHPOQT 1920
 QY 1921 LHPGNIPESSQONLLOPLKPSPESSDNLVSAFTSDGALSVPSLSAPQGTSTVTVGATV 1980
 Db 1921 LHPGNIPESSQONLLOPLKPSPESSDNLVSAFTSDGALSVPSLSAPQGTSTVTVGATV 1980
 QY 1981 NSQAQAQOPAMTSSRYGTFTDHLKLVDMVARDAMNLSGHRGSKGHNNYSGPMARFES 2040
 Db 1981 NSQAQAQOPAMTSSRYGTFTDHLKLVDMVARDAMNLSGHRGSKGHNNYSGPMARFES 2040
 QY 2041 APGOLCISMTSNLGSAPISASATSLGHFTKSCPPQOYGFPAFPFGAOMSGTGAPAPQ 2100
 Db 2041 APGOLCISMTSNLGSAPISASATSLGHFTKSCPPQOYGFPAFPFGAOMSGTGAPAPQ 2100
 QY 2101 PLGQFOPVGTASLQNFNLSNLOKSIINPFGSNLRTT 2136
 Db 2101 PLGQFOPVGTASLQNFNLSNLOKSIINPFGSNLRTT 2136

RESULT 2


```

Db      840 PSSGRHGRTRTKHRYKSVASRSRHEKTSRPKILINWKNKGRVVECOLETHNRKMYT 899
Qy      901 FKFDLDGDNPEELIATIMVNDFFILAIERESFVDQVREIIIEKADEMLSEDSVEEGDGL 960
Db      900 FKFDLDGDNPEELIATIMVNDFFILAIERESFVDQVREIIIEKADEMLSEDSVEEGDGL 959
Qy      961 ESLOGKDYGSGSGCKEGERKOPIPASSMPOQIGIPSSILQVHVHSGREFIVSPVES 1020
Db      960 ESLOGKDYGSGSGCKEGERKOPIPASSMPOQIGIPSSILQVHVHSGREFIVSPVES 1019
Qy      1021 RLRSKVPSEITDTVAASTAQSPGMNLSHSASISLQOAFSELRRAQMTGPTAPNPF 1080
Db      1020 RLRSKVPSEITDTVAASTAQSPGMNLSHSASISLQOAFSELRRAQMTGPTAPNPF 1079
Qy      1081 SHTGPTFPVVPFPLSSLAGVPTTAATAPVATSPENDISTSVIQSEVTVPTBEGLAGV 1140
Db      1080 SHTGPTFPVVPFPLSSLAGVPTTAATAPVATSPENDISTSVIQSEVTVPTBEGLAGV 1139
Qy      1141 ATSTGVTSAGGLPIPVSESPVLSVSGSITIPAVVSISSPSLOVPTSTSEIVSSTA 1200
Db      1140 ATSTGVTSAGGLPIPVSESPVLSVSGSITIPAVVSISSPSLOVPTSTSEIVSSTA 1199
Qy      1201 LPSVTVSATSASAGSITATGPKPPAVVSOQAAGSTTVGATLTSVSTTSFPSTASQLS 1260
Db      1200 LPSVTVSATSASAGSITATGPKPPAVVSOQAAGSTTVGATLTSVSTTSFPSTASQLS 1259
Qy      1261 IQLSSSTSTPPLAEVTVVSAHSLDKTSHSSTTGCLAFSLASPSSSSPGAGVSIYSGPGG 1320
Db      1260 IQLSSSTSTPPLAEVTVVSAHSLDKTSHSSTTGCLAFSLASPSSSSPGAGVSIYSGPGG 1319
Qy      1321 LHPVIVPEVIASTPLIPQAAGPTSTPLIPQVSIPIVQVPAVPAVQGLTIHSQOPAL 1380
Db      1320 LHPVIVPEVIASTPLIPQAAGPTSTPLIPQVSIPIVQVPAVPAVQGLTIHSQOPAL 1379
Qy      1381 LPPNPHTHCEPVDSDTOPKAFGIDIDIKLEKLSLFEHSSSGAQHSAVSELESLVIES 1440
Db      1380 LPPNPHTHCEPVDSDTOPKAFGIDIDIKLEKLSLFEHSSSGAQHSAVSELESLVIES 1439
Qy      1441 TVTGGIPTTANAPSKLSTSTSTCLPPTNMLPLGYALPVTVVTVPGQVSTVSTTSQVYK 1500
Db      1440 TVTGGIPTTANAPSKLSTSTSTCLPPTNMLPLGYALPVTVVTVPGQVSTVSTTSQVYK 1499
Qy      1501 PGTAPSKRPLTKAPLPIVTELPAGTLPSBQLPPPGSLQSQOQPLEDLAOLRRTLSP 1560
Db      1500 PGTAPSKRPLTKAPLPIVTELPAGTLPSBQLPPPGSLQSQOQPLEDLAOLRRTLSP 1559
Qy      1561 EXITVTSVAVGSMAAPTAITEAGTOPQKGVSYQVKEGVLATSSGAGVFEMKGRFQVSVA 1620
Db      1560 EXITVTSVAVGSMAAPTAITEAGTOPQKGVSYQVKEGVLATSSGAGVFEMKGRFQVSVA 1619
Qy      1621 DGAQKEGKNKEDAKSVHFESESTSSSVLSSSSPSTLYVKEPENGITIPGISSDVPEBSAH 1680
Db      1620 DGAQKEGKNKEDAKSVHFESESTSSSVLSSSSPSTLYVKEPENGITIPGISSDVPEBSAH 1679
Qy      1681 KTTASEAKSDTGOPTKVGKRFVTTTANKVGRFSYKTEKDTITDPTKESGPVSPPTMDEQ 1740
Db      1680 KTTASEAKSDTGOPTKVGKRFVTTTANKVGRFSYKTEKDTITDPTKESGPVSPPTMDEQ 1739
Qy      1741 AVLPVAVIPKKEKPELSEPSHNLNGPSDEEAFLSKDVDDGSGSPHSPHQLSSKSLPSQNTL 1800
Db      1740 AVLPVAVIPKKEKPELSEPSHNLNGPSDEEAFLSKDVDDGSGSPHSPHQLSSKSLPSQNTL 1799
Qy      1801 SOSLSNNSNNSYMSDNDSDIEDDELKLELRRLRKHKEIODLOSQKHIEISLYTTLG 1860
Db      1800 SOSLSNNSNNSYMSDNDSDIEDDELKLELRRLRKHKEIODLOSQKHIEISLYTTLG 1859
Qy      1861 KVPAPAVIIPAPAPISGRRRPTKSGKSSSSSLGKNSPOLSGTLSSQSAASYVHPQOT 1920
Db      1860 KVPAPAVIIPAPAPISGRRRPTKSGKSSSSSLGKNSPOLSGTLSSQSAASYVHPQOT 1919
Qy      1921 LHPGNTIPESGONLQPLKPSPSNDNLYSAFTSDGAISVPSLSAPGQSTSTVTCATV 1980

```

```

Db      1920 LHPGNTIPESGONLQPLKPSPSNDNLYSAFTSDGAISVPSLSAPGQSTSTVTCATV 1979
Qy      1981 NSQAQAQOPPAMTSSRKGTFTDDIHLKLVDMWARDANMLSGRRSGKGMNVEGPMARKXS 2040
Db      1980 NSQAQAQOPPAMTSSRKGTFTDDIHLKLVDMWARDANMLSGRRSGKGMNVEGPMARKXS 2039
Qy      2041 APGQLCTSMTSNLSGSAPIASASATSLGHTFKSCPPQOYGFPAATPFQAOMSGTGPAPO 2100
Db      2040 APGQLCTSMTSNLSGSAPIASASATSLGHTFKSCPPQOYGFPAATPFQAOMSGTGPAPO 2099
Qy      2101 PLGQFQPVGTASLQNFNINSLQKSIINPPGSLNRTT 2136
Db      2100 PLGQFQPVGTASLQNFNINSLQKSIINPPGSLNRTT 2135

RESULT 3
ABO44384
ID ABO44384 standard; protein; 2229 AA.
XX
AC ABO44384;
XX
DT 26-SEP-2003 (first entry)
XX
DE Novel human protein kinase #4.
XX
KW Human; kinase; enzyme; cosmetic application; nutraceutical application.
XX
OS Homo sapiens.
XX
PN US6541252-B1.
XX
PD 01-APR-2003.
XX
PF 14-MAY-2001; 2001US-00854856.
XX
PR 19-MAY-2000; 2000US-0206015P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Hilbun E, Donoho G, Turner CA;
XX
DR WPI; 2003-575927/54.
XX
DR N-PSDB; ACH03767.
XX
PT New nucleic acid encoding novel human proteins, useful in cosmetic and
PS nutraceutical applications.
XX
PS Disclosure; Page; 11pp; English.
XX
CC The invention relates to a new isolated nucleic acid encoding a novel
CC human protein kinase. The nucleic acid is useful in cosmetic and
CC nutraceutical applications. The present sequence represents the amino
CC acid sequence of a novel human protein kinase. Note: The sequence data
CC for this patent did not from part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=6541252B1
XX
SQ Sequence 2229 AA.

Query Match 99.5%; Score 10753.5; DB 7; Length 2229;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2136; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

Qy      1 MSGGAARKQSTPSSLFSPAPAPKXGSSSDSVGKLGAAADATYGRTEFRRRRHT 60
Db      1 MSGGAARKQSTPSSLFSPAPAPKXGSSSDSVGKLGAAADATYGRTEFRRRRHT 60
Qy      61 MDKSRGAAATTTTTHRRFRFRSVICSNATALELPGPLSLPPQSIIPAAVQSAPEEPH 120
Db      61 MDKSRGAAATTTTTHRRFRFRSVICSNATALELPGPLSLPPQSIIPAAVQSAPEEPH 120
Qy      121 REETVTATATSOVAQOPAAAPGQAVAGPAPSTVSTSKDPVVSQPSLVSGKEEPP 180

```

Db 121 REEVTATATISQVAAQPPAAAAAPEGQAVAGAPSTVPSISKDRPVSPSLVSGKEPPP 180
 QY 181 ARSSGGGSAKEPOEERSQOQDDIELEETKAVGMSNDRFLKFDIEIGRSGFKTVYKGLD 240
 Db 181 ARSSGGGSAKEPOEERSQOQDDIELEETKAVGMSNDRFLKFDIEIGRSGFKTVYKGLD 240
 QY 241 TETTVAVWCELODRKLTSEORRPFKEAEMKGLQHPNIVRFYDSWESTVYKGCICVLV 300
 Db 241 TETTVAVWCELODRKLTSEORRPFKEAEMKGLQHPNIVRFYDSWESTVYKGCICVLV 300
 QY 301 TELMTSGTLKTYLRFKVMKIKVLRSWCRQILKGLQPLHTTPPIIHRDLKCDNIFITGP 360
 Db 301 TELMTSGTLKTYLRFKVMKIKVLRSWCRQILKGLQPLHTTPPIIHRDLKCDNIFITGP 360
 QY 361 TGSVYKIGDLGLATLKRAFSFAKSVIGTEPFNAPEMVEEYKDESUVVAFGMCMLEMTSEY 420
 Db 361 TGSVYKIGDLGLATLKRAFSFAKSVIGTEPFNAPEMVEEYKDESUVVAFGMCMLEMTSEY 420
 QY 421 PYSECQNAAOIYRRVTSQVSPASFDKVAIPVKEIIIEGCIQRONDERYSIKOLLNHAFFQ 480
 Db 421 PYSECQNAAOIYRRVTSQVSPASFDKVAIPVKEIIIEGCIQRONDERYSIKOLLNHAFFQ 480
 QY 481 EETVRVELAEEDGEXIAIKMLRIDIKKLKGKYDNDAIESPDLERDVEDVAQEM 540
 Db 481 EETVRVELAEEDGEXIAIKMLRIDIKKLKGKYDNDAIESPDLERDVEDVAQEM 540
 QY 541 VESGYCEGDHKTAKAIKDRVSLIKRRREGROLVREEOKKOEESLXQOVBQSSASQ 600
 Db 541 VESGYCEGDHKTAKAIKDRVSLIKRRREGROLVREEOKKOEESLXQOVBQSSASQ 600
 QY 601 TGKOLPASTGIPFTASTASVSTQVPEPEADHQOIQYQOPSISVLSDGTVDSDGQ 660
 Db 601 TGKOLPASTGIPFTASTASVSTQVPEPEADHQOIQYQOPSISVLSDGTVDSDGQ 660
 QY 661 SSYTERSVSSQOQVSVSGQHEQAHSNGTVPGHPISTVQAQSQCHGYPPSSVAQSGSQ 720
 Db 661 SSYTERSVSSQOQVSVSGQHEQAHSNGTVPGHPISTVQAQSQCHGYPPSSVAQSGSQ 720
 QY 721 QPSSSSLTGVSSSQPIQHPQOQOQIQQTAAPQOQVQYSLSTQTSSEATTAQAPSPQAP 780
 Db 721 QPSSSSLTGVSSSQPIQHPQOQOQIQQTAAPQOQVQYSLSTQTSSEATTAQAPSPQAP 780
 QY 781 QVLPQVSAHQ----- 791
 Db 781 QVLPQVSAHQ----- 791
 QY 792 -----STQGVQVAPAEVAV 807
 Db 792 -----STQGVQVAPAEVAV 807
 QY 841 VNOPIYSNLLVPMGVGVGQOVQVSPGGSIAQAPTTSSQOAVLESTQGVQVAPAEVAV 900
 Db 841 VNOPIYSNLLVPMGVGVGQOVQVSPGGSIAQAPTTSSQOAVLESTQGVQVAPAEVAV 900
 QY 808 AOPQATOPPTLASVDSAHSDVASGMSDGENVPSSSGRHEGRTTKAHYKRSRSH 867
 Db 808 AOPQATOPPTLASVDSAHSDVASGMSDGENVPSSSGRHEGRTTKAHYKRSRSH 867
 QY 901 AOPQATOPPTLASVDSAHSDVASGMSDGENVPSSSGRHEGRTTKAHYKRSRSH 960
 Db 901 AOPQATOPPTLASVDSAHSDVASGMSDGENVPSSSGRHEGRTTKAHYKRSRSH 960
 QY 868 KTSRPKRLIIVNSKGRVVECOLETHNRKMTFKFDLDGDNPEIATIMWNNDFILAI 927
 Db 868 KTSRPKRLIIVNSKGRVVECOLETHNRKMTFKFDLDGDNPEIATIMWNNDFILAI 927
 QY 961 KTSRPKRLIIVNSKGRVVECOLETHNRKMTFKFDLDGDNPEIATIMWNNDFILAI 1020
 Db 961 KTSRPKRLIIVNSKGRVVECOLETHNRKMTFKFDLDGDNPEIATIMWNNDFILAI 1020
 QY 928 RESFVDVREIIEKADMLSEVSVPEPSOGLESLOKQDYGFSQKLEGFKOPIPA 987
 Db 928 RESFVDVREIIEKADMLSEVSVPEPSOGLESLOKQDYGFSQKLEGFKOPIPA 987
 QY 1021 RESFVDVREIIEKADMLSEVSVPEPSOGLESLOKQDYGFSQKLEGFKOPIPA 1080
 Db 1021 RESFVDVREIIEKADMLSEVSVPEPSOGLESLOKQDYGFSQKLEGFKOPIPA 1080
 QY 988 SSMPOOIGIPTSSLTQVHSAGRFIVSPVESRLRSKVPSEITDTVAASAQSPGM 1047
 Db 988 SSMPOOIGIPTSSLTQVHSAGRFIVSPVESRLRSKVPSEITDTVAASAQSPGM 1047
 QY 1081 SSMPOOIGIPTSSLTQVHSAGRFIVSPVESRLRSKVPSEITDTVAASAQSPGM 1140
 Db 1081 SSMPOOIGIPTSSLTQVHSAGRFIVSPVESRLRSKVPSEITDTVAASAQSPGM 1140
 QY 1048 LSHSASLSLQAFSELRRAQMTGENTAPNFSHTGPTFPVVPPLISIAAGVPTTAAT 1107
 Db 1048 LSHSASLSLQAFSELRRAQMTGENTAPNFSHTGPTFPVVPPLISIAAGVPTTAAT 1107
 QY 1141 LSHSASLSLQAFSELRRAQMTGENTAPNFSHTGPTFPVVPPLISIAAGVPTTAAT 1200
 Db 1141 LSHSASLSLQAFSELRRAQMTGENTAPNFSHTGPTFPVVPPLISIAAGVPTTAAT 1200
 QY 1108 APVPATSSPNDISTSVIOSEVTVPEEGIGAGVASTGVVTSGLPIPVSESPVLSVV 1167
 Db 1108 APVPATSSPNDISTSVIOSEVTVPEEGIGAGVASTGVVTSGLPIPVSESPVLSVV 1167
 QY 1201 APVPATSSPNDISTSVIOSEVTVPEEGIGAGVASTGVVTSGLPIPVSESPVLSVV 1260
 Db 1201 APVPATSSPNDISTSVIOSEVTVPEEGIGAGVASTGVVTSGLPIPVSESPVLSVV 1260

QY 1168 SSITIPAVVSIISTSPLOVPTSTSEIIVSSTALYPSTVATSASAGSGTATPGKPPA 1227
 Db 1261 SSITIPAVVSIISTSPLOVPTSTSEIIVSSTALYPSTVATSASAGSGTATPGKPPA 1320
 QY 1228 VVSQQAAGSTTVGATLTSVSTTSPSTASQSLSTQSSSTSTPTLAEVTVVSAHSLDKTS 1287
 Db 1321 VVSQQAAGSTTVGATLTSVSTTSPSTASQSLSTQSSSTSTPTLAEVTVVSAHSLDKTS 1380
 QY 1288 HSSSTGLAFSLASAPSSSSSGAGVSYISQPGCHLPIVPSVASTPTLPOAGPTSTPL 1347
 Db 1381 HSSSTGLAFSLASAPSSSSSGAGVSYISQPGCHLPIVPSVASTPTLPOAGPTSTPL 1440
 QY 1348 LPQVPSIPPLVOPANVPAVOQTILHSQOPALBNQPHNHCPEVSDTOPKAGIDIDIK 1407
 Db 1441 LPQVPSIPPLVOPANVPAVOQTILHSQOPALBNQPHNHCPEVSDTOPKAGIDIDIK 1500
 QY 1408 TLEBKRLSFESEHSSGAQHASVSLSTSVIESTVTPGIPPTTAVAPSKLITSTSTCLP 1467
 Db 1501 TLEBKRLSFESEHSSGAQHASVSLSTSVIESTVTPGIPPTTAVAPSKLITSTSTCLP 1560
 QY 1468 TNLPLGTVALPVTVPVTPPGQVSTPVSTTSGVPRGTAPSKRPLTKAPVLEVGTLEPAGTL 1527
 Db 1561 TNLPLGTVALPVTVPVTPPGQVSTPVSTTSGVPRGTAPSKRPLTKAPVLEVGTLEPAGTL 1620
 QY 1528 PSEQLPFPQPSLTQSQOPLLEDIAQRLTSEEXIIVTSAVGVSWAAPPTATTEAGTOP 1587
 Db 1621 PSEQLPFPQPSLTQSQOPLLEDIAQRLTSEEXIIVTSAVGVSWAAPPTATTEAGTOP 1680
 QY 1588 QKGVSYKEGVPVLTSSGAGVFKMGPRQVSAADGAQKEGKNSEDAKSHFESSSTESS 1647
 Db 1681 QKGVSYKEGVPVLTSSGAGVFKMGPRQVSAADGAQKEGKNSEDAKSHFESSSTESS 1740
 QY 1648 VLSSSSPRESLTVRPEPNGITIPGISDVPESAKHTTASEAKSDTQPTKYGRFOVTTAN 1707
 Db 1741 VLSSSSPRESLTVRPEPNGITIPGISDVPESAKHTTASEAKSDTQPTKYGRFOVTTAN 1800
 QY 1708 KYGRFVSXTKEDKITDTPKKEGPVAPSPFMOLDEAVIPAVIPKKEKPELSPSHNGSSSD 1767
 Db 1801 KYGRFVSXTKEDKITDTPKKEGPVAPSPFMOLDEAVIPAVIPKKEKPELSPSHNGSSSD 1860
 QY 1768 PEAAFLSRDVEDGSGSPHSPHOLSSKSLPSQNLSSQSLNSFNSSYMSSDNESEDIEEDLK 1827
 Db 1861 PEAAFLSRDVEDGSGSPHSPHOLSSKSLPSQNLSSQSLNSFNSSYMSSDNESEDIEEDLK 1920
 QY 1828 LELRLRLDKLKEIOLQOSQKHEISLYTKLQKVPNAVILPPAAPLSGRRRPTYSKGS 1887
 Db 1921 LELRLRLDKLKEIOLQOSQKHEISLYTKLQKVPNAVILPPAAPLSGRRRPTYSKGS 1980
 QY 1888 KSSRSSSLGKNSPOLSGNLSSQSAASVLPQOQLHPNGNIPESGONQLOPLKPSPSDN 1947
 Db 1981 KSSRSSSLGKNSPOLSGNLSSQSAASVLPQOQLHPNGNIPESGONQLOPLKPSPSDN 2040
 QY 1948 LYSAFSTDGAIIVPSLSAPQGTSTNTVGAIVNSQAAQAPAMTSSRKGTFTDDLHLK 2007
 Db 2041 LYSAFSTDGAIIVPSLSAPQGTSTNTVGAIVNSQAAQAPAMTSSRKGTFTDDLHLK 2100
 QY 2068 GHFTKMCPPQOQYGFATPFGAQMSGTGPAPQPLQFOFQVGVASLQNFISLQOSISN 2127
 Db 2161 GHFTKMCPPQOQYGFATPFGAQMSGTGPAPQPLQFOFQVGVASLQNFISLQOSISN 2220
 QY 2128 PGGSNIRTT 2136
 Db 2221 PGGSNIRTT 2229

RESULT 4
 ABP71619
 ID ABP71619 standard; protein; 2382 AA.

XX AC ABP71619;
 XX DT 29-MAY-2003 (first entry)
 XX DE Human MNK1 protein.
 XX NM MNK1, MNK4, hypotensive; gene therapy; antisense therapy; vaccine;
 XX KW pseudohypoadosteronism type II; PHA II; human.
 XX OS Homo sapiens.
 XX PN W02003007793-A2.
 XX PD 30-JAN-2003.
 XX PF 17-JUL-2002; 2002W0-US022744.
 XX PR 17-JUL-2001; 2001US-0306084P.
 XX (UYVA) UNIV YALE.
 XX (LIFT) LIFTON R P.
 XX (WILS) WILSON F H.
 XX (CHOA) CHOATE K.
 XX (NELS) NELSON-WILLIAMS C.
 XX (ISHI) ISHIKAWA K.
 XX LIFTON RP, Wilson FH, Choate K, Nelson-Williams C, Ishikawa K;
 XX WPI; 2003-289755/28.
 XX DR N-PSDB; ABZ59199.
 XX PT Mammalian MNK nucleic acids and proteins useful for the diagnosis,
 XX PT prevention and treatment of hypertension and pseudohypoadosteronism
 XX type II.
 XX PS Claim 13; Page 101-109; 116pp; English.
 XX CC The invention relates to novel nucleic acids encoding mammalian MNK,
 XX CC especially human MNK1 and MNK4 proteins. The MNK polypeptides can be
 XX CC expressed by standard recombinant methodology. The MNK polynucleotides,
 XX CC polypeptides and modulators may be used in the prevention, diagnosis and
 XX CC treatment of diseases associated with inappropriate MNK expression such
 XX CC as hypertension and pseudohypoadosteronism type II (PHA II). The
 XX CC present sequence represents the human MNK1 protein
 XX SQ Sequence 2382 AA;
 Query Match 98.8%; Score 10677; DB 6; Length 2382;
 Best Local Similarity 89.6%; Pred. No. 0;
 Matches 2135; Conservative 0; Mismatches 1; Indels 246; Gaps 1;

DB 301 TELMTSGTLKTYLKRFRKMKIKVLRSMCRQILKGLQFLHTRTPPIIHRDLKCNIFITGP 360
 QY 361 TGSVKIGDLGLATLTKASPAKSVYIGTPEPMAPMEYEKIDESVDVYAFGCMEMATSEY 420
 DB 361 TGSVKIGDLGLATLTKASPAKSVYIGTPEPMAPMEYEKIDESVDVYAFGCMEMATSEY 420
 QY 421 PYSECNNAQIYRRVMSGVKPAFEDKVALPEVEIIEGCRONKDEYSIKDLINHAFFQ 480
 DB 421 PYSECNNAQIYRRVMSGVKPAFEDKVALPEVEIIEGCRONKDEYSIKDLINHAFFQ 480
 QY 481 EETGVVELAEEDDGEKIAIKMLRIEDIKLKGYKONDAIEFSPDLERDVEDVAQEM 540
 DB 481 EETGVVELAEEDDGEKIAIKMLRIEDIKLKGYKONDAIEFSPDLERDVEDVAQEM 540
 QY 541 VESGYVCEGDHKTMAKAIKDRVSLIKRQRQULVREBEKKQEBESLKQVQESSASQ 600
 DB 541 VESGYVCEGDHKTMAKAIKDRVSLIKRQRQULVREBEKKQEBESLKQVQESSASQ 600
 QY 601 TGIKOLPSASTGIPASTTSASVSTOVEPEPEADQHQLOYOQPSISVLDGTVDSGQ 660
 DB 601 TGIKOLPSASTGIPASTTSASVSTOVEPEPEADQHQLOYOQPSISVLDGTVDSGQ 660
 QY 661 SSVFESRVSSQQTVSYSQHEQAHSSTGVPGHIPSTVQAQSPHGYPPSSVAQGSQ 720
 DB 661 SSVFESRVSSQQTVSYSQHEQAHSSTGVPGHIPSTVQAQSPHGYPPSSVAQGSQ 720
 QY 721 QPSSSILTVSSSQPIQHPQOQOGIQTAPPOQTQVYSLSQTSSTSSATTAQPPSQAP 780
 DB 721 QPSSSILTVSSSQPIQHPQOQOGIQTAPPOQTQVYSLSQTSSTSSATTAQPPSQAP 780
 QY 781 QVLPOVSAGKQ----- 791
 DB 781 QVLPOVSAGKQVLVSPQVPTIQEBPQIPVATQPSVSVPHSGAHFLPVQGLPPLPQY 840
 QY 792 ----- 791
 DB 841 VSGIPISTPHVSTAQNTGSSLPITMAAGITQPLITLASATTAIPGVTIVBSQPLTL 900
 QY 792 ----- 791
 DB 901 QVYTLQESQVHPOLLQPAVSMGIPANLGAAEVPLSSGVLYOGPPRLPEQYGDNSI 960
 QY 792 ----- 791
 DB 961 APSNNVASVCIHSTVLSPPMPTEVLAIPGYFPITVQYVESNLIVPMGIVGQVQVSPG 1020
 QY 792 -----STQVGSQVADAEPAVAQAQATQPTTLASSVDSASHSDVASGMS 834
 DB 1021 GSIAQAPFTSSQQAVESTQGVSGVAPAEPAVAQAQATQPTTLASSVDSASHSDVASGMS 1080
 QY 835 DGENVNVSSSGRHEGRTTKRHYRVSRSRHEKTSPKRIILNVSKGRVVECOLETH 894
 DB 1081 DGENVNVSSSGRHEGRTTKRHYRVSRSRHEKTSPKRIILNVSKGRVVECOLETH 1140
 QY 895 NRKAVTFKFDLDGDNPEIATIMVNNDFILAIERESFVDOVREILIERADENLSEDSVER 954
 DB 1141 NRKAVTFKFDLDGDNPEIATIMVNNDFILAIERESFVDOVREILIERADENLSEDSVER 1200
 QY 955 EGDGLSLOKDDYGFSGSQKLEGEKOPIPASSMPOQIGIPFSSILTOYVHSGRFTV 1014
 DB 1201 EGDGLSLOKDDYGFSGSQKLEGEKOPIPASSMPOQIGIPFSSILTOYVHSGRFTV 1260
 QY 1015 SPVPESLRRESKVPESITITVAASTAQSGMNLSSHASSLSLQQAASELRAQMTGPN 1074
 DB 1261 SPVPESLRRESKVPESITITVAASTAQSGMNLSSHASSLSLQQAASELRAQMTGPN 1320
 QY 1075 TAPPNFSHTGTFPVVPPFLSLIAGVPTTAATAPVATSSPPNDISTSVIQSEVYVTE 1134
 DB 1321 TAPPNFSHTGTFPVVPPFLSLIAGVPTTAATAPVATSSPPNDISTSVIQSEVYVTE 1380
 QY 1135 EGIAGVASTGVVTSGLPIPVSESPVLSVSSITIPAVVSIITSSPSIQVPTISEI 1194

Db 1381 EGIAGVATSTGVTWVSGGLPIPPVSESPVLSSVSVSTITIPAVVSTSTTSPISQVPTSTSEI 1440
 QY 1195 VVSSVATLVSVYVTSATSSAGSGSTATPGKPPAVVSOQAAGSTTGATLTSTSTTSPS 1254
 Db 1441 VVSSVATLVSVYVTSATSSAGSGSTATPGKPPAVVSOQAAGSTTGATLTSTSTTSPS 1500
 QY 1255 TASQSLSTLSSSTSTPTLAEVTVVSAHSLDKTSHSSTTGALFSLAPSSSSPGAGVSSY 1314
 Db 1501 TASQSLSTLSSSTSTPTLAEVTVVSAHSLDKTSHSSTTGALFSLAPSSSSPGAGVSSY 1560
 QY 1315 ISQPGHPLVIPSIVTASTPTLLPQAAGPTSTPLPQVPSIPPLVQPVANVAVOQTLLHS 1374
 Db 1561 ISQPGHPLVIPSIVTASTPTLLPQAAGPTSTPLPQVPSIPPLVQPVANVAVOQTLLHS 1620
 QY 1375 QPOBALPNOPTHCEVSDTOPKAPGIDIKTLEELKRLSFSEHSSGAGHVSLET 1434
 Db 1621 QPOBALPNOPTHCEVSDTOPKAPGIDIKTLEELKRLSFSEHSSGAGHVSLET 1680
 QY 1435 SLVIESVTPGIPPTTAAVAPSKLTSTSTCLPPTNLPLGTVALPVTVPVTPGOVSTPVST 1494
 Db 1681 SLVIESVTPGIPPTTAAVAPSKLTSTSTCLPPTNLPLGTVALPVTVPVTPGOVSTPVST 1740
 QY 1495 TTSGVKGTAASKPPLTKAPVLPVGTLPACTLPSBOLPFPFGSILTOSQOPLDLDLQOL 1554
 Db 1741 TTSGVKGTAASKPPLTKAPVLPVGTLPACTLPSBOLPFPFGSILTOSQOPLDLDLQOL 1800
 QY 1555 RRTLSPEXITVTSVAVGPVMAAPATAITEAGTQPOKGVQVEGCVLATSSAGVFFKMGKF 1614
 Db 1801 RRTLSPEXITVTSVAVGPVMAAPATAITEAGTQPOKGVQVEGCVLATSSAGVFFKMGKF 1860
 QY 1615 QVSVAAADGAQKEGKNSKEDAKSVHEBSTSSSVLSSSSPSTLVKEPENGITTPGISSD 1674
 Db 1861 QVSVAAADGAQKEGKNSKEDAKSVHEBSTSSSVLSSSSPSTLVKEPENGITTPGISSD 1920
 QY 1675 VPESAHKTTAESAASDQGPPTKVGPFQVTTANKVGPSPVSKTDDKTITDTKEGPVAPSP 1734
 Db 1921 VPESAHKTTAESAASDQGPPTKVGPFQVTTANKVGPSPVSKTDDKTITDTKEGPVAPSP 1980
 QY 1735 FMDLEQAVLPVAVIPKKEPELSEPSHLNGPSSDPEAAFLSDVDGSGSPHSPHQLSKS 1794
 Db 1981 FMDLEQAVLPVAVIPKKEPELSEPSHLNGPSSDPEAAFLSDVDGSGSPHSPHQLSKS 2040
 QY 1795 LPSQNLQSLSNSFNSSSYMSSDNEDIEDEDLKLRLRDLKHLKEIQDLQSRQKHIES 1854
 Db 2041 LPSQNLQSLSNSFNSSSYMSSDNEDIEDEDLKLRLRDLKHLKEIQDLQSRQKHIES 2100
 QY 1855 LYTTLGKVPVPAVILIPPAAPLSGRRRRPTKSGKSSSSSSISLGNKSPQLSGNLSCQSAASY 1914
 Db 2101 LYTTLGKVPVPAVILIPPAAPLSGRRRRPTKSGKSSSSSSISLGNKSPQLSGNLSCQSAASY 2160
 QY 1915 LHPQQTLPHPGNIPESSQNLQPLKSPSSSDNLYSAFTSPDGAISVPSLSAPGGTSTTN 1974
 Db 2161 LHPQQTLPHPGNIPESSQNLQPLKSPSSSDNLYSAFTSPDGAISVPSLSAPGGTSTTN 2220
 QY 1975 TVGATVNSQAQAOPPAMTSSRKGTFTDDLHKLVDNMAPDAMNLSGRGSGKHNNYEGPG 2034
 Db 2221 TVGATVNSQAQAOPPAMTSSRKGTFTDDLHKLVDNMAPDAMNLSGRGSGKHNNYEGPG 2280
 QY 2035 MARFESAPOULCISMTSNLGGSAPIISAASATSLGHFTKSMCPPOQYFPATPFAQMSGT 2094
 Db 2281 MARFESAPOULCISMTSNLGGSAPIISAASATSLGHFTKSMCPPOQYFPATPFAQMSGT 2340
 QY 2095 GGPAQPLGQFQPVGTASLQNFNISNLQKSIINPPGSLNRTT 2136
 Db 2341 GGPAQPLGQFQPVGTASLQNFNISNLQKSIINPPGSLNRTT 2382

DT 26-SEP-2003 (first entry)
 XX
 DE Novel human protein kinase #1.
 XX
 KM Human; kinase; enzyme; cosmetic application; nutraceutical application.
 XX
 OS Homo sapiens.
 XX
 PN US6541252-B1.
 XX
 PD 01-APR-2003.
 XX
 PF 14-MAY-2001; 2001US-00854856.
 XX
 PR 19-MAY-2000; 2000US-0206015P.
 XX
 PA (TEXT-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hilbun E, Donoho G, Turner CA;
 XX
 PI WPI: 2003-575927/54.
 DR N-PSDB; ACH03764.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutraceutical applications.
 XX
 PS Disclosure; Page: 11pp; English.
 XX
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents the amino
 CC acid sequence of a novel human protein kinase. Note: The sequence data
 CC for this patent did not form part of the printed specification but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=6541252B1
 CC
 XX
 SQ Sequence 2382 AA;
 Query Match 98.8%; Score 10677; DB 7; Length 2382;
 Best Local Similarity 89.7%; Pred. No. 0;
 Matches 2136; Conservative 0; Mismatches 0; Indels 246; Gaps 1;
 QY 1 MSGGAARKQSGTSGSLFLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTTEYRRRHT 60
 Db 1 MSGGAARKQSGTSGSLFLSPAPAPKNGSSSDSVGEKLGAAADAVTGRTTEYRRRHT 60
 QY 61 MDXDSRGAAATTTTTEHFRFRRSVYCDSNATALELPGLPLSLPQSPIDPAVPGSAPPEPH 120
 Db 61 MDXDSRGAAATTTTTEHFRFRRSVYCDSNATALELPGLPLSLPQSPIDPAVPGSAPPEPH 120
 QY 121 REETVATATSOVAQCPAAAPAEQVAVAPBPTVSSSTSKORPVQSPISVSGKEPP 180
 Db 121 REETVATATSOVAQCPAAAPAEQVAVAPBPTVSSSTSKORPVQSPISVSGKEPP 180
 QY 181 ARSGGGGSAKPEDEERSQOQDDIELEETAVAGSNDGRFLKFDIEIGRGSFKTVYKGLD 240
 Db 181 ARSGGGGSAKPEDEERSQOQDDIELEETAVAGSNDGRFLKFDIEIGRGSFKTVYKGLD 240
 QY 241 TETTVEAVACHELDRKLTKESEORFKEEAEMLKGLQHPNIVRFYDSWESTVKGKCTIVL 300
 Db 241 TETTVEAVACHELDRKLTKESEORFKEEAEMLKGLQHPNIVRFYDSWESTVKGKCTIVL 300
 QY 301 TELMTSGTLKTYLKRFFVMKIKVLRSCROILNGLOLHRTPTPIIHRDLKCNIFITGP 360
 Db 301 TELMTSGTLKTYLKRFFVMKIKVLRSCROILNGLOLHRTPTPIIHRDLKCNIFITGP 360
 QY 361 TGSVKIGDLGLATLKRASFASKSVIGTPEFMAPEMEYKEDSVYVAFGCMLEMATSEY 420
 Db 361 TGSVKIGDLGLATLKRASFASKSVIGTPEFMAPEMEYKEDSVYVAFGCMLEMATSEY 420
 QY 421 PYSECQMAAQIYRRVYSGVPASFDKVAIEVKEIIEGCIKONKDERYSIKDLINHAFFQ 480
 Db 421 PYSECQMAAQIYRRVYSGVPASFDKVAIEVKEIIEGCIKONKDERYSIKDLINHAFFQ 480

RESULT 5
 ABO44381
 ID ABO44381 standard; protein; 2382 AA.
 XX
 AC ABO44381;
 XX

01-APR-2003.
 14-MAY-2001; 2001US-00854856.
 19-MAY-2000; 2000US-0206015P.
 (LEXI-) LEXICON GENETICS INC.
 Walke DW, Hilbun E, Donoho G, Turner CA;
 WPI; 2003-575927/54.
 N-PSDB; ACH03778.
 New nucleic acid encoding novel human proteins, useful in cosmetic and
 nutritional applications.
 Disclosure; Page; 11pp; English.
 The invention relates to a new isolated nucleic acid encoding a novel
 human protein kinase. The nucleic acid is useful in cosmetic and
 nutritional applications. The present sequence represents the amino
 acid sequence of a novel human protein kinase. Note: The sequence data
 for this patent did not form part of the printed specification but was
 obtained in electronic format directly from USPTO at
 seqdata.uspto.gov/sequence.html?docid=6541252B1
 Sequence 2108 AA;
 Query Match 98.5%; Score 10646; DB 7; Length 2108;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

601 TGIKLPASASTGIPTASTTSASVSTOVEPEPEADHQOLQYQPSISVLSDGTVDGSG 660
 661 SSVFTESSVSSQOCVTSVSGQHEQAHSTGTGPHGIPSTVQAOSQPHGYPPSSVQAQSGSG 720
 661 SSVFTESSVSSQOCVTSVSGQHEQAHSTGTGPHGIPSTVQAOSQPHGYPPSSV----- 713
 721 QPSSSLTGVSASQPIQHPOQOQGIQOTAPPOQTVOVSLSGTSTSTSTTAPQPSOPAP 780
 714 -----QSGIQOTAPPOQTVOVSLSGTSTSTSTTAPQPSOPAP 752
 781 QVLPQVSAKOSTQGVSAQVAPAEVAVAOQATPPTLASSVDSAHSDVASMGDGENV 840
 753 QVLPQVSAKOSTQGVSAQVAPAEVAVAOQATPPTLASSVDSAHSDVASMGDGENV 812
 841 PSSSGRHEGRTTKRHKYKSVSRSRHEKTSRPKRIINVSNGKDRVVECOLETHNRKMT 900
 813 PSSSGRHEGRTTKRHKYKSVSRSRHEKTSRPKRIINVSNGKDRVVECOLETHNRKMT 872
 901 FKFDLDGNEBEIATIMVNNDFILATERESFVDQVREIIRKADMLSEDSVVEEGDQGL 960
 873 FKFDLDGNEBEIATIMVNNDFILATERESFVDQVREIIRKADMLSEDSVVEEGDQGL 932
 961 ESLQKQDYFGSGQKLEGEFFKQPIPASSMPQOIGIPTSSLTVYHSAGRFIVSPYES 1020
 933 ESLQKQDYFGSGQKLEGEFFKQPIPASSMPQOIGIPTSSLTVYHSAGRFIVSPYES 992
 1021 RLRESKVFSEIITDYVAASTAQSFGMNLSHSASLSIQAFSELRRAQOMTEGPTADPNF 1080
 993 RLRESKVFSEIITDYVAASTAQSFGMNLSHSASLSIQAFSELRRAQOMTEGPTADPNF 1052
 1081 SHTGPTFPVVPPLSIAGVPTTAATAAPAPATSSPNDISTSVTQSEVVPVEEGTAGV 1140
 1053 SHTGPTFPVVPPLSIAGVPTTAATAAPAPATSSPNDISTSVTQSEVVPVEEGTAGV 1112
 1141 ATSTGVTSGGLPIPVSESPVLSVVSITIPAVASISTSPSLQVPTSTSEIIVSSTA 1200
 1113 ATSTGVTSGGLPIPVSESPVLSVVSITIPAVASISTSPSLQVPTSTSEIIVSSTA 1172
 1201 LPSVTVSATSASAGSSTAPGPKPAPVAVSQAAAGSTTVACATLTSVTTSPSTASQLS 1260
 1173 LPSVTVSATSASAGSSTAPGPKPAPVAVSQAAAGSTTVACATLTSVTTSPSTASQLS 1232
 1261 IQSSSTPTTAEIYVAVSHSLDKTSHSSTTGLASLSAPSSSSSGAGVSSYISQPG 1320
 1233 IQSSSTPTTAEIYVAVSHSLDKTSHSSTTGLASLSAPSSSSSGAGVSSYISQPG 1292
 1321 LHPVLVPSVIASTPILPOAAGPTSTPLLPQVPSIPPLVQVAVNAVVOQLLHISQOPAL 1380
 1293 LHPVLVPSVIASTPILPOAAGPTSTPLLPQVPSIPPLVQVAVNAVVOQLLHISQOPAL 1352
 1381 LHPQPHTHCPENDSDTPKAPGIDDKITLBEKRSLSFSRSHSSGAQHAASYLSTLVIES 1440
 1353 LHPQPHTHCPENDSDTPKAPGIDDKITLBEKRSLSFSRSHSSGAQHAASYLSTLVIES 1412
 1441 TVTPGIPPTAVASKLLTSTSTCLPPTNLPLGTVALPTPVYTPQOVSPVSTTSYGVK 1500
 1413 TVTPGIPPTAVASKLLTSTSTCLPPTNLPLGTVALPTPVYTPQOVSPVSTTSYGVK 1472
 1501 PGTAPSKPPLTKAPVLPVGTETLPAGTLPSBOLPPFPGPSILTOSQQLLELDLDAQLRRTLS 1560
 1473 PGTAPSKPPLTKAPVLPVGTETLPAGTLPSBOLPPFPGPSILTOSQQLLELDLDAQLRRTLS 1532
 1533 EXITTVSAVGPVMAAPTAITEAGTPOKGVSVKGPVLAUSSGAGVFMGRFOVSVA 1592
 1561 EXITTVSAVGPVMAAPTAITEAGTPOKGVSVKGPVLAUSSGAGVFMGRFOVSVA 1620
 1621 DGAQKGGKKSBEAKVHFESSSTSSSVLSSSPSTLVYPEBNGATTIPISDVESAH 1680
 1593 DGAQKGGKKSBEAKVHFESSSTSSSVLSSSPSTLVYPEBNGATTIPISDVESAH 1652
 1681 KTTASAKSDTQOPTVGRFOVTTTANKVGRFVSSTEDKIITDKKEGPVASDPFMDLEQ 1740

Db 1653 KTTA5EAKSDTQGTXYGRFQVTTANKGRF5SKTEDIKITDTKEGEPVAPPEFMDLEQ 1712

Qy 1741 AVLPAYIPKKEKPELSPSHLNGPSSPPEAFPLSRDVEDSGSPHSHQISSKSLPEQNTL 1800

Db 1713 AVLPAYIPKKEKPELSPSHLNGPSSPPEAFPLSRDVEDSGSPHSHQISSKSLPEQNTL 1772

Qy 1801 SQTLSNFSNYSMSDNEDEDEDLKLELRDLKLEKEIODEQSQKHEIBSLYTKLG 1860

Db 1773 SQTLSNFSNYSMSDNEDEDEDLKLELRDLKLEKEIODEQSQKHEIBSLYTKLG 1832

Qy 1861 KVPAPVILPPAPPLSGRRRRTKSGKSSRSSSLGNKSPQLSGNLSGQSAAYLHPQQT 1920

Db 1833 KVPAPVILPPAPPLSGRRRRTKSGKSSRSSSLGNKSPQLSGNLSGQSAAYLHPQQT 1892

Qy 1921 LHPPGNIPESGQNLLOPLKPSPESSDNLKSAFTSDGAI5VPSLSAPQGGSSNTTVATV 1980

Db 1893 LHPPGNIPESGQNLLOPLKPSPESSDNLKSAFTSDGAI5VPSLSAPQGGSSNTTVATV 1952

Qy 1981 NSQAAQAPPAMTSSRGFTFDLHKLVDNWARDAMNLSGRSGSKGMNYEGPQMARKFS 2040

Db 1953 NSQAAQAPPAMTSSRGFTFDLHKLVDNWARDAMNLSGRSGSKGMNYEGPQMARKFS 2012

Qy 2041 APGQLCSMTSNLGGSAPI5AASATSLGHFTKSNCPQYQGFPAFPFGAQMSTGCPAPQ 2100

Db 2013 APGQLCSMTSNLGGSAPI5AASATSLGHFTKSNCPQYQGFPAFPFGAQMSTGCPAPQ 2072

Qy 2101 PLGQFPVGTASLQENFISNLQKSI5NPPGSLNFTT 2136

Db 2073 PLGQFPVGTASLQENFISNLQKSI5NPPGSLNFTT 2108

RESULT 7

ADCG9049

ID ADCG9049 standard; protein: 2108 AA.

XX

AC ADCG9049;

XX

DT 01-JAN-2004 (first entry)

XX

DE Human KRP protein - SEQ ID 2.

XX

KM anti-HIV; antiallergic; antiinflammatory; antianaemic; antiparkinsonian;

KM nootropic; anticonvulsant; antiarteriosclerotic; antiaschematic;

KM immunosuppressive; antihypertoid; cytosstatic; hepatotropic; dermatological;

KM antidiabetic; nephrotropic; antistout; thyromimetic; neuroprotective;

KM osteopathic; antiatherosclerotic; antiparasitic; antihelminthic; antibacterial;

KM uteropathic; ophthalmological; antirheumatic; haemostatic; antiproliferative;

KM virucide; protozoacide; fungicide; kinase; phosphatase; KRP;

KM cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;

KM cancer; developmental; mental retardation; neurological;

KM Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;

KM diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;

KM helminthic infection; transgenic; gene therapy; human; enzyme.

XX

OS Homo sapiens.

XX

PN W02003033680-AA.

XX

PD 24-APR-2003.

XX

PF 17-OCT-2002; 2002W0-US033723.

XX

PR 19-OCT-2001; 2001US-0345474P.

PR 02-NOV-2001; 2001US-0343910P.

PR 13-NOV-2001; 2001US-0333098P.

PR 16-NOV-2001; 2001US-0332424P.

PR 30-NOV-2001; 2001US-0334288P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Bandman O, Baughn MR, Becha SD, Borowsky ML, Dugan BM;

PI Emerling BM, Forsythe IU, Gandhi AR, Gorvad AE, Griffin JA;

PI Gururajan R, Hafalia AJA, Khan FA, Lal PG, Lee EA, Lee SY;

PI Lindquist EA, Lu DM, Lu Y, Marguis JP, Nguyen DB, Ariznu CS;

PI Ramkumar J, Recipon SA, Richardson TW, Swarnakar A, Tang YT;

PI Thornton MB, Iran UK, Chawla NK, Warren BA, Yang U, Yao MG, Yue H;

PI Zebatyadian Y;

DR WPI: 2003-403214/38.

DR N-PSDB; ADC99101.

XX

PT New human kinases and phosphatases and polynucleotides, useful for

PT diagnosing, treating or preventing autoimmune or inflammatory disorders

PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,

PT cancer or hepatitis.

XX

PS Claim 1; SEQ ID NO 2; 424bp; English.

XX

CC The invention relates to a novel isolated polypeptide which is a human

CC kinase and phosphatase (KRP). The KRP polypeptides, polynucleotides,

CC agonists and antagonists are useful for diagnosing, treating or

CC preventing cell proliferative disorders such as atherosclerosis,

CC cirrhosis, hepatitis and cancer, developmental disorders e.g. mental

CC retardation, neurological disorders including Alzheimer's disease and

CC Parkinson's disease, autoimmune and inflammatory disorders such as

CC Crohn's disease and diabetes mellitus and finally, viral, bacterial,

CC fungal, parasitic, protozoan or helminthic infections. Furthermore, the

CC polynucleotides encoding KRP may be useful for creating transgenic

CC animals to model human disease, as well as during gene therapy

CC procedures. The current sequence is that of the human KRP protein of the

CC invention.

XX

XX

Sequence 2108 AA;

Query Match 98.5%; Score 10646; DB 7; Length 2108;

Best Local Similarity 98.6%; Pred. No. 0;

Matches 2107; Conservative 0; Mismatches 1; Indels 28; Gaps 1;

Qy 1 MSGAAEKGSTPGSLPLSPAPAPKXSSSDSVGEKLGAAADAVTGRTEYRRRHT 60

Db 1 MSGAAEKGSTPGSLPLSPAPAPKXSSSDSVGEKLGAAADAVTGRTEYRRRHT 60

Qy 61 MDKDSRGAAATTTTTERFRFRSVTC5NNTALELPLPLSLPQPSIPAAVPSAPPEPH 120

Db 61 MDKDSRGAAATTTTTERFRFRSVTC5NNTALELPLPLSLPQPSIPAAVPSAPPEPH 120

Qy 121 REETVTATNSOYAAQOPPAAPAEQAVAPAPSTVSSSTKDPVQPSLVGSKPEPP 180

Db 121 REETVTATNSOYAAQOPPAAPAEQAVAPAPSTVSSSTKDPVQPSLVGSKPEPP 180

Qy 181 ARSGGGGSAKEPQEBRSQOQDDIELETRAVGMSNDGRELKPDIEIGSGFKTVYGLD 240

Db 181 ARSGGGGSAKEPQEBRSQOQDDIELETRAVGMSNDGRELKPDIEIGSGFKTVYGLD 240

Qy 241 TETTVEAVMCELODRKLTKEERQFKKEAEMLGLQHPNIVREFDSWESYVKGKCCIVL 300

Db 241 TETTVEAVMCELODRKLTKEERQFKKEAEMLGLQHPNIVREFDSWESYVKGKCCIVL 300

Qy 301 TELMTSGTLKTYLKRFPVMKIKYLR5WCROILKGLQFLHRTPEIHRDLKCDNIFITGP 360

Db 301 TELMTSGTLKTYLKRFPVMKIKYLR5WCROILKGLQFLHRTPEIHRDLKCDNIFITGP 360

Qy 361 TGSVKIGDLGLATLKRA5PAKSVIGTPEFMAPEMYEKEYDESVDVYAFGCMLEMATSEY 420

Db 361 TGSVKIGDLGLATLKRA5PAKSVIGTPEFMAPEMYEKEYDESVDVYAFGCMLEMATSEY 420

Qy 421 PYSECQNAAOIYRRVYSGVAPASDKAIPEVVEITIGCRONNDEY5TKDLINHAFFQ 480

Db 421 PYSECQNAAOIYRRVYSGVAPASDKAIPEVVEITIGCRONNDEY5TKDLINHAFFQ 480

Qy 481 EETGVRELAEEDDGEKIAIKMLRIDIKLKQKYDNALIESFLERDVEDVAQEM 540

Db 481 EETGVRELAEEDDGEKIAIKMLRIDIKLKQKYDNALIESFLERDVEDVAQEM 540

Qy 541 VESGYVCGDHKTMAKIKDVSIIKRRQROLVREDEKKQESSLKOQVQSSASQ 600

Db 541 VESGVCEGDHKTMAKAIKDRVSLIKRRKQRLVREOEKKOEESLKQOVQSSASQ 600
 QY 601 TGKQLPFASTGIPASTTASVSTQVEPEPEADQHOOLQYQOQPSIVLSLSDGTVDSSQG 660
 Db 601 TGIQQLPASTGIPASTTASVSTQVEPEPEADQHOOLQYQOQPSIVLSLSDGTVDSSQG 660
 QY 661 SSVPFESRVSSQOQTVSYSGHQAHSSTGTVPQGHIPSTVQAQSQPHGVYPSSVVAQOQSQG 720
 Db 661 SSVPFESRVSSQOQTVSYSGHQAHSSTGTVPQGHIPSTVQAQSQPHGVYPSSVVAQOQSQG 720
 QY 721 QPSSSLTGVSSSQIHPQOQOQGIQOQAPPOQTVQYSLSQTSSTSEKTTQAPVSQOPAP 780
 Db 721 QPSSSLTGVSSSQIHPQOQOQGIQOQAPPOQTVQYSLSQTSSTSEKTTQAPVSQOPAP 780
 QY 781 QVLPOVASGOSTQGVSOVAPAEPAVAVQOPATQPTTASVDSASHSVASGMDGNV 840
 Db 753 QVLPOVASGOSTQGVSOVAPAEPAVAVQOPATQPTTASVDSASHSVASGMDGNV 840
 QY 841 PSSSGRHEGRTTKHRYKRSVRSRHEKTSRPKLILNVSNKGRVVECOLETHNRKMT 900
 Db 841 PSSSGRHEGRTTKHRYKRSVRSRHEKTSRPKLILNVSNKGRVVECOLETHNRKMT 900
 QY 901 FKFLDGNPEEIAITIMNDFILAIERESFVDQREIIEKADMLSDVSVPEBGDGL 960
 Db 873 FKFLDGNPEEIAITIMNDFILAIERESFVDQREIIEKADMLSDVSVPEBGDGL 960
 QY 961 ESLOGKDDYGSQKLEGEFKOPIPASMPQOIGIPSSLTQVHSGRFFIYSPVES 1020
 Db 933 ESLOGKDDYGSQKLEGEFKOPIPASMPQOIGIPSSLTQVHSGRFFIYSPVES 992
 QY 1021 RLRESKVPSEITDTVAASQSPQMLSHSASSLSLOAFSELRRAQMTGEPNAPNF 1080
 Db 993 RLRESKVPSEITDTVAASQSPQMLSHSASSLSLOAFSELRRAQMTGEPNAPNF 1052
 QY 1081 SHTGPTPVVPPFLSSINGVTTAAATAPVPATSPPPNDISTVYQSEVYTPTEBGLAGV 1140
 Db 1053 SHTGPTPVVPPFLSSINGVTTAAATAPVPATSPPPNDISTVYQSEVYTPTEBGLAGV 1112
 QY 1141 ATSTGVTSVGGGLPIPPVESPVLSGVSSITIPAVVSISSTPSPLOVFTSSEIIVSSSTA 1200
 Db 1113 ATSTGVTSVGGGLPIPPVESPVLSGVSSITIPAVVSISSTPSPLOVFTSSEIIVSSSTA 1172
 QY 1201 LYPSTVATSASAGSJTATPGKPPAVVSOQAAGSTTVGATLTVSFTTSPSTASQLS 1260
 Db 1173 LYPSTVATSASAGSJTATPGKPPAVVSOQAAGSTTVGATLTVSFTTSPSTASQLS 1232
 QY 1261 IQLSSSTPTLAEVTVVSAHSLDKTSHSSTTGLAFSLAPSSSSSPGAGVSYISQPGG 1320
 Db 1233 IQLSSSTPTLAEVTVVSAHSLDKTSHSSTTGLAFSLAPSSSSSPGAGVSYISQPGG 1292
 QY 1321 LHPVTPSVIASTPTLPQAGPTSTPLPOVPSIPPLVQPVANVAVQOTLIHSQOPAL 1380
 Db 1293 LHPVTPSVIASTPTLPQAGPTSTPLPOVPSIPPLVQPVANVAVQOTLIHSQOPAL 1352
 QY 1381 LPNQPHTHCPEVDSPTQKAPGIDIDIKTLEKRLSLFSEHSSSGAQAHSVLSLETSLVIES 1440
 Db 1353 LPNQPHTHCPEVDSPTQKAPGIDIDIKTLEKRLSLFSEHSSSGAQAHSVLSLETSLVIES 1412
 QY 1441 TVTPTGIPITTAAPSKLTLSTTCLPPTNLPLGTALVTVTVTPVPGVSTVSTTSGVK 1500
 Db 1413 TVTPTGIPITTAAPSKLTLSTTCLPPTNLPLGTALVTVTVTPVPGVSTVSTTSGVK 1472
 QY 1501 PGTAPSKRPLTKAPVLPVGTLPAGTLPSBQLPPPGSLTQSQQLPDLDLQAKRLTISP 1560
 Db 1473 PGTAPSKRPLTKAPVLPVGTLPAGTLPSBQLPPPGSLTQSQQLPDLDLQAKRLTISP 1532
 QY 1561 EXITVTSVAVGVSMAPTAITEAGTQPKGVQVVEGVLTATSSAGVFKMGRFQVSVAA 1620
 Db 1533 EXITVTSVAVGVSMAPTAITEAGTQPKGVQVVEGVLTATSSAGVFKMGRFQVSVAA 1592
 QY 1621 DGAQEGKXKEDASVHFESSTSSSVLSSSSPSTLVKPEPNGITTPGISSDVPEBAH 1680
 Db 1593 DGAQEGKXKEDASVHFESSTSSSVLSSSSPSTLVKPEPNGITTPGISSDVPEBAH 1652

QY 1681 KTTASEAKSDPTGQPTKVGREQVTTANKVGRFSVSKTEDKITDTPKKGFPVAPPPMDLEQ 1740
 Db 1653 KTTASEAKSDPTGQPTKVGREQVTTANKVGRFSVSKTEDKITDTPKKGFPVAPPPMDLEQ 1712
 QY 1741 AVLPVAPPKKEKPELSPSHLNGSPSPDEAFLSRVDVDSGSHSPHQLSSKSLPSQNL 1800
 Db 1713 AVLPVAPPKKEKPELSPSHLNGSPSPDEAFLSRVDVDSGSHSPHQLSSKSLPSQNL 1772
 QY 1801 SOSLSNFSNYSWSDNESDIEDLKLRLRLDKLKEIQLDLSRQKHEIESLYTKLG 1860
 Db 1773 SOSLSNFSNYSWSDNESDIEDLKLRLRLDKLKEIQLDLSRQKHEIESLYTKLG 1832
 QY 1861 KVPVAVIIPPAAPLSGRRRRTKSGKSSSRSSSLGNKSPOLSGNLSGQSAASVLIHPQOT 1920
 Db 1833 KVPVAVIIPPAAPLSGRRRRTKSGKSSSRSSSLGNKSPOLSGNLSGQSAASVLIHPQOT 1892
 QY 1921 LHPGNIPESGONLLOPLKPSPSDNLVSAFTSDGAI SVPSLSAPGQGISSTNTVGATV 1980
 Db 1893 LHPGNIPESGONLLOPLKPSPSDNLVSAFTSDGAI SVPSLSAPGQGISSTNTVGATV 1952
 QY 1981 NSQAQAOPPAMTSRRKGTFTDIDLKLVDMWAPAMNLGSRGSKGHMNYEGPGMARKFS 2040
 Db 1953 NSQAQAOPPAMTSRRKGTFTDIDLKLVDMWAPAMNLGSRGSKGHMNYEGPGMARKFS 2012
 QY 2041 APGOLCISMTSNLGSAPISASATSLGHFTKSCPCPOQYGFPAATPFACMSGTGSPAPQ 2100
 Db 2013 APGOLCISMTSNLGSAPISASATSLGHFTKSCPCPOQYGFPAATPFACMSGTGSPAPQ 2072
 QY 2101 PLGQFQPVGTASLQNFNINISNLOKISINPPGNSLRTT 2136
 Db 2073 PLGQFQPVGTASLQNFNINISNLOKISINPPGNSLRTT 2108
 RESULT 8
 ABO4392 standard; protein; 2201 AA.
 ID ABO4392 standard; protein; 2201 AA.
 AC ABO4392;
 XX
 DT 26-SEP-2003 (first entry)
 XX
 DE Novel human protein kinase #12.
 XX
 KW Human; kinase; enzyme; cosmetic application; nutraceutical application.
 XX
 OS Homo sapiens.
 XX
 PN US6541252-B1.
 XX
 PD 01-APR-2003.
 XX
 PF 14-MAY-2001; 2001US-00854856.
 XX
 PR 19-MAY-2000; 2000US-0206015P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Hilbun E, Donoho G, Turner CA;
 DR WPI; 2003-575927/54.
 XX
 DR N-PSDB; ACH03775.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutraceutical applications.
 XX
 PS Disclosure; Page; 11pp; English.
 XX
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents the amino
 CC acid sequence of a novel human protein kinase. Note: The sequence data
 CC for this patent did not from part of the printed specification but was

CC obtained in electronic format directly from USPRO at
 CC seqdata.uspto.gov/sequence.html?docid=6541252B1
 XX
 SQ Sequence 2201 AA;

Query Match 97.9%; Score 10589.5; DB 7; Length 2201;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 0; Indels 121; Gaps 2;

QY 1 MSGAAEKOSSTPGSLFLSPAPAPKNGSSSDSSVGEKLGAAADAVVGRTEYRRRHT 60
 Db 1 MSGAAEKOSSTPGSLFLSPAPAPKNGSSSDSSVGEKLGAAADAVVGRTEYRRRHT 60
 QY 61 MDKSRGAATTTTTEHRRFRSVICDSNATALEPLGLPLSPQSPIDPAVPOSAPPEPH 120
 Db 61 MDKSRGAATTTTTEHRRFRSVICDSNATALEPLGLPLSPQSPIDPAVPOSAPPEPH 120
 QY 121 REETVTAATQVAAQPPAAAAGPQAVAGPAPSTVPSSTSKDRVSOPLVSGKEEPP 180
 Db 121 REETVTAATQVAAQPPAAAAGPQAVAGPAPSTVPSSTSKDRVSOPLVSGKEEPP 180
 QY 181 ARSGSGGSAKEPQERSSQODDIELETKAVGMSNDRFLKFDIEIGRSGFTVYKGLD 240
 Db 181 ARSGSGGSAKEPQERSSQODDIELETKAVGMSNDRFLKFDIEIGRSGFTVYKGLD 240
 QY 241 TETTVAVAMELODKLTKESEKFEKAEMLKGLQHPNIVRFYDSWESTVKGKCIYLV 300
 Db 241 TETTVAVAMELODKLTKESEKFEKAEMLKGLQHPNIVRFYDSWESTVKGKCIYLV 300
 QY 301 TELMTSGTLTKYKFKVMKIKVLRSCROILKGLQPLHTPTPIIHRDLKCDNIFITGP 360
 Db 301 TELMTSGTLTKYKFKVMKIKVLRSCROILKGLQPLHTPTPIIHRDLKCDNIFITGP 360
 QY 361 TGSVYIGDLGATLKRAFAKSVGTPEFMAPEWEEKYDESVMYARGMCLMATESEY 420
 Db 361 TGSVYIGDLGATLKRAFAKSVGTPEFMAPEWEEKYDESVMYARGMCLMATESEY 420
 QY 421 PYSECQNAQAQYRRVTSVGPASFPDKVAIPYKKEIIIEGICIRONKDERYSIKDLINHAFFQ 480
 Db 421 PYSECQNAQAQYRRVTSVGPASFPDKVAIPYKKEIIIEGICIRONKDERYSIKDLINHAFFQ 480
 QY 481 EETGVAVELAEEDGEEKIAIKMLRIEDIKLKKGKYNONEAIBESFEDLERVPEVDAEM 540
 Db 481 EETGVAVELAEEDGEEKIAIKMLRIEDIKLKKGKYNONEAIBESFEDLERVPEVDAEM 540
 QY 541 VESGVCEGDEKHTAKAIKDRVSLIKRKROOLVREOEKKOEESLKQOVEOSSASQ 600
 Db 541 VESGVCEGDEKHTAKAIKDRVSLIKRKROOLVREOEKKOEESLKQOVEOSSASQ 600
 QY 601 TGIKQLPASGTIPTASTTSASVTOVPEPEADQHQLOQYQOQPSISVLSDGTVDSSQG 660
 Db 601 TGIKQLPASGTIPTASTTSASVTOVPEPEADQHQLOQYQOQPSISVLSDGTVDSSQG 660
 QY 661 SSVTETESVSSQOQVSYSGOHEQAHSTGTVPGHPTVQAOQOPHGVPPSSVAGQSG 720
 Db 661 SSVTETESVSSQOQVSYSGOHEQAHSTGTVPGHPTVQAOQOPHGVPPSSVAGQSG 720
 QY 721 QPSSSILTVSSSQPIQHPQOQOQIOQAPPOQVQVSLQSTSTSSSEATTQAPVSOQAP 780
 Db 721 QPSSSILTVSSSQPIQHPQOQOQIOQAPPOQVQVSLQSTSTSSSEATTQAPVSOQAP 780
 QY 781 QVLPOVSAGKQ-----QGGIQQTAPOQVQVSLQSTSTSSSEATTQAPVSOQAP 791
 Db 781 QVLPOVSAGKQ-----QGGIQQTAPOQVQVSLQSTSTSSSEATTQAPVSOQAP 791
 QY 792 -----STQGVSOVAPAEVAV 807
 Db 792 -----STQGVSOVAPAEVAV 807
 QY 808 AOPQATOPTTLASVDSHSDVASGMSDGENENPVSSSGRHEGRITTKHYRVSRSRHE 867
 Db 808 AOPQATOPTTLASVDSHSDVASGMSDGENENPVSSSGRHEGRITTKHYRVSRSRHE 867
 QY 873 AOPQATOPTTLASVDSHSDVASGMSDGENENPVSSSGRHEGRITTKHYRVSRSRHE 932
 Db 873 AOPQATOPTTLASVDSHSDVASGMSDGENENPVSSSGRHEGRITTKHYRVSRSRHE 932

QY 868 KTSRPKRLILNVSNKGDVEVECOLETHNRKMVTFEFDLDGDNPEEIAITMNNDFILAE 927
 Db 933 KTSRPKRLILNVSNKGDVEVECOLETHNRKMVTFEFDLDGDNPEEIAITMNNDFILAE 992
 QY 928 RESFVDQVREIIEKADEMLSEDDVSVEPEQDQLESLQCKDDYGFSGQKLEGEFKQPIPA 987
 Db 993 RESFVDQVREIIEKADEMLSEDDVSVEPEQDQLESLQCKDDYGFSGQKLEGEFKQPIPA 1052
 QY 988 SSMPOQIGITPTSSILTOVHSHGRRTIVSPVESRLREKVPSEITTDVVAASTAOSPQGN 1047
 Db 1053 SSMPOQIGITPTSSILTOVHSHGRRTIVSPVESRLREKVPSEITTDVVAASTAOSPQGN 1112
 QY 1048 LSHSASSLSLQOASSELRAQMTGEPNTAPNFSHTGTFFVPPFELSIAIGVPTTAAT 1107
 Db 1113 LSHSASSLSLQOASSELRAQMTGEPNTAPNFSHTGTFFVPPFELSIAIGVPTTAAT 1172
 QY 1108 APVPATSSPPNDISTSVIQSEVTVPTBEGLAGVATSTGVTSGLPIPPVBSPLYSSV 1167
 Db 1173 APVPATSSPPNDISTSVIQSEVTVPTBEGLAGVATSTGVTSGLPIPPVBSPLYSSV 1232
 QY 1168 SSITIPAVVSIITSPSLQVPTSEIYVSTALYPSVTVSATASAGGATATGPKRPA 1227
 Db 1233 SSITIPAVVSIITSPSLQVPTSEIYVSTALYPSVTVSATASAGGATATGPKRPA 1292
 QY 1228 VVSQOAGSTTVGATLISVSTTSPFSTAQSLQLSSTSTPLLAETVVSASHSDKTS 1287
 Db 1293 VVSQOAGSTTVGATLISVSTTSPFSTAQSLQLSSTSTPLLAETVVSASHSDKTS 1352
 QY 1288 HSSTTGLAFSLAPSSSSSPGAGVSSYSQPGIHLPIVPSVIASTPILPQAAGPTSTPL 1347
 Db 1353 HSSTTGLAFSLAPSSSSSPGAGVSSYSQPGIHLPIVPSVIASTPILPQAAGPTSTPL 1412
 QY 1348 LPOVPSIPPLVQAPVAPVQOOLIHQOPQALLPNOPTHCPVDSQTKAGIDIK 1407
 Db 1413 LPOVPSIPPLVQAPVAPVQOOLIHQOPQALLPNOPTHCPVDSQTKAGIDIK 1472
 QY 1408 TLEBKRLSLESEHSSGASHASVLETSVLESTVTFGTPTTAAVPSKLTSTSTCLP 1467
 Db 1473 TLEBKRLSLESEHSSGASHASVLETSVLESTVTFGTPTTAAVPSKLTSTSTCLP 1532
 QY 1468 TNLPLGTVALPVPVTPGVSTVSTTSGVKGTAAPSKPPLTKAPVLPVGTBLPAGTL 1527
 Db 1533 TNLPLGTVALPVPVTPGVSTVSTTSGVKGTAAPSKPPLTKAPVLPVGTBLPAGTL 1592
 QY 1528 PSEOLPPPGPSLQSOQPLEDLAOLRRLTSPKXITVTSVAVGVSAAPALTEAGQP 1587
 Db 1593 PSEOLPPPGPSLQSOQPLEDLAOLRRLTSPKXITVTSVAVGVSAAPALTEAGQP 1652
 QY 1588 QKGVSYQKEGVFLATSSGAGVFKMGREFQVSVAAADGAKGEGNKSEDAKSVHEESTSSS 1647
 Db 1653 QKGVSYQKEGVFLATSSGAGVFKMGREFQVSVAAADGAKGEGNKSEDAKSVHEESTSSS 1712
 QY 1648 VLSSSPESTLVKEPENGTIPGISDVPSAHKTTTSEAKSDTGOPTKVGREFVYTTAN 1707
 Db 1713 VLSSSPESTLVKEPENGTIPGISDVPSAHKTTTSEAKSDTGOPTKVGREFVYTTAN 1772
 QY 1708 KVGFRSVSKTEDKTTDKKGGPVASPPFMDLEQAVLPAVLPKKEKPELSEESHNGPSSD 1767
 Db 1773 KVGFRSVSKTEDKTTDKKGGPVASPPFMDLEQAVLPAVLPKKEKPELSEESHNGPSSD 1832
 QY 1768 PEAFLSRDYDDGSGSPHSPHQLSSKSLPSQNSQSLSNSNSNSYMSDNESEDLEDELK 1827
 Db 1833 PEAFLSRDYDDGSGSPHSPHQLSSKSLPSQNSQSLSNSNSNSYMSDNESEDLEDELK 1892
 QY 1828 LELRLRDXHLKEIODIOSRQKEHIESLYTKLGAVPPAVIIPPAAPLSGRRRRPTKSGS 1867
 Db 1893 LELRLRDXHLKEIODIOSRQKEHIESLYTKLGAVPPAVIIPPAAPLSGRRRRPTKSGS 1952
 QY 1868 KSSRSSLSLGNKSPOLSGNLSGQSAAYLHPQOOLHPGNTPEGQONOLLOPLKSPSSDN 1947
 Db 1953 KSSRSSLSLGNKSPOLSGNLSGQSAAYLHPQOOLHPGNTPEGQONOLLOPLKSPSSDN 2012
 QY 1948 LYSAFSTDGAISVPSLAPGQSTSTVTVGATVNSQAQAQPPAMTSSRKGTFTDIDLHL 2007

Db 2013 LYSATSDGAISSVSLASAPGGGISTNTVGATVNSQAQAQPPAMTSSRKCTFTDDHLKL 2072
 Qy 2008 VDNARDDAMNLSGRSGSKGHNNVGGPGMARKPSAPGOLCISMTSNLGGASAPISASATSL 2067
 Db 2073 VDNARDDAMNLSGRSGSKGHNNVGGPGMARKPSAPGOLCISMTSNLGGASAPISASATSL 2132
 Qy 2068 GHFTKSMCPPOQYGFPAFPFAGQWSGTGGAPQPLGQFPQVGTASLQNFNISNLQKSIISN 2127
 Db 2133 GHFTKSMCPPOQYGFPAFPFAGQWSGTGGAPQPLGQFPQVGTASLQNFNISNLQKSIISN 2132
 Qy 2128 PPGSNLRTT 2136
 Db 2193 PPGSNLRTT 2201
 RESULT 9
 ABO44389
 ID ABO44389 standard; protein, 2354 AA.
 AC ABO44389;
 XX 26-SEP-2003 (first entry)
 DT 26-SEP-2003 (first entry)
 XX
 DE Novel human protein kinase #9.
 KW Human; kinase; enzyme; cosmetic application; nutraceutical application.
 OS Homo sapiens.
 XX
 PN US6541252-B1.
 PD 01-APR-2003.
 XX
 PF 14-MAY-2001; 2001US-00854856.
 XX
 PR 19-MAY-2000; 2000US-0206015P.
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 FI Walke DW, Hilbun E, Donoho G, Turner CA;
 DR WPI; 2003-575927/54.
 DR N-PSDB; ACH03772.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutraceutical applications.
 PS
 PS Disclosure; Page; 11pp; English.
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents the amino
 CC acid sequence of a novel human protein kinase. Note: The sequence data
 CC for this patent did not from part of the printed specification but was
 CC obtained in electronic format directly from USPTO at
 CC segdata.uspto.gov/sequence.html?docid=6541252B1
 CC
 SQ Sequence 2354 AA;
 Query Match 97.2%; Score 10513; DB 7; Length 2354;
 Best Local Similarity 88.5%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 0; Indels 274; Gaps 2;

Db 121 REETVTATATSOVAQPPAAPGQAVAGAPSTVPSSTSKDPVSPQSLVGSKEPP 180
 Qy 181 ARSGGGGSAKEPEEESQOQDDIELEETAYVAGSNNGRFLKPIETIGRGSFKTYKGLD 240
 Db 181 ARSGGGGSAKEPEEESQOQDDIELEETAYVAGSNNGRFLKPIETIGRGSFKTYKGLD 240
 Qy 241 TETVEVAMCELODRKLTTSERQRFKEAEMLKGLQHPNIVRFDSWSETVKGGKCIYLV 300
 Db 241 TETVEVAMCELODRKLTTSERQRFKEAEMLKGLQHPNIVRFDSWSETVKGGKCIYLV 300
 Qy 301 TELMTSGTLTKYLRFRVWKIKVLRSWCRQILKGLQPLHTRTPPIIRHDLKCDNIPTIGP 360
 Db 301 TELMTSGTLTKYLRFRVWKIKVLRSWCRQILKGLQPLHTRTPPIIRHDLKCDNIPTIGP 360
 Qy 361 TGSVKIGDIGLATLKRASFASVGTPEPEMAPEMYEEKYDESVDVYAFGCMLEMAVSEY 420
 Db 361 TGSVKIGDIGLATLKRASFASVGTPEPEMAPEMYEEKYDESVDVYAFGCMLEMAVSEY 420
 Qy 421 PYSECQNAAQIYRVTSGVGPASFDKYAIPVKEIIIEGCIIRONDERYSIKDLINHAFFQ 480
 Db 421 PYSECQNAAQIYRVTSGVGPASFDKYAIPVKEIIIEGCIIRONDERYSIKDLINHAFFQ 480
 Qy 481 EETGVAVELAEEDDEGEKIAIKLMIRIDIKLKKYKQNEAIEFSFDLERDVPBDVQEM 540
 Db 481 EETGVAVELAEEDDEGEKIAIKLMIRIDIKLKKYKQNEAIEFSFDLERDVPBDVQEM 540
 Qy 541 VESGVCEGDHKTAKAIIKORVSLIKRREOROLVREBOEKKQESSLKQVROSSASQ 600
 Db 541 VESGVCEGDHKTAKAIIKORVSLIKRREOROLVREBOEKKQESSLKQVROSSASQ 600
 Qy 601 TGIKQLPSASTGILPTASTTSASVSTOVEPEPEADQHQQLQYQPSISVLSDGTVDSSQG 660
 Db 601 TGIKQLPSASTGILPTASTTSASVSTOVEPEPEADQHQQLQYQPSISVLSDGTVDSSQG 660
 Qy 661 SSVFTESRVSSQQTVSQSGHEQASHSTGTVPGHIPSTVQAQSGPHGVPPSSVAQSGSQG 720
 Db 661 SSVFTESRVSSQQTVSQSGHEQASHSTGTVPGHIPSTVQAQSGPHGVPPSSVAQSGSQG 720
 Qy 721 QPSSSSILTVSSSQPIQHQQOQGIQOTAPRQCVQVYSLQSTSSSEATTQAPVQAP 780
 Db 721 QPSSSSILTVSSSQPIQHQQOQGIQOTAPRQCVQVYSLQSTSSSEATTQAPVQAP 780
 Qy 781 QVLPQVSAQKQ- 791
 Db 781 QVLPQVSAQKQ- 791
 Qy 791 QVLPQVSAQKQ- 791
 Db 791 QVLPQVSAQKQ- 791
 Qy 813 VSQIPISIPHVSTAQTGFSLSLPTMAAGITQPLTLTASSATTAAIPGVSTVPSQLPTLL 872
 Db 813 VSQIPISIPHVSTAQTGFSLSLPTMAAGITQPLTLTASSATTAAIPGVSTVPSQLPTLL 872
 Qy 792 791
 Db 792 791
 Qy 873 QPVTQPSQVHPOLLQPAVQSMGIPANLQAAEVLSSGDVLYQGFPRLPPQYIGDSNTI 932
 Db 873 QPVTQPSQVHPOLLQPAVQSMGIPANLQAAEVLSSGDVLYQGFPRLPPQYIGDSNTI 932
 Qy 792 791
 Db 792 791
 Qy 933 APSGNVASVCHSTVLAHPMPTEVLAITPGFPTVQVYVESNLLVPMGVGGVQVQVQPG 992
 Db 933 APSGNVASVCHSTVLAHPMPTEVLAITPGFPTVQVYVESNLLVPMGVGGVQVQVQPG 992
 Qy 792 834
 Db 792 834
 Qy 993 GSILAQPTTSSQQAULESTQGVSVAPAPVAVQPAQTPPTTLASSVDSASHSVASMS 1052
 Db 993 GSILAQPTTSSQQAULESTQGVSVAPAPVAVQPAQTPPTTLASSVDSASHSVASMS 1052
 Qy 835 DGENNVPESSGRHGRGRTTKRHYRYSVSRSHKTSRKLILNVNSNGDAVVEGQLETH 894
 Db 835 DGENNVPESSGRHGRGRTTKRHYRYSVSRSHKTSRKLILNVNSNGDAVVEGQLETH 894
 Qy 1053 DGENNVPESSGRHGRGRTTKRHYRYSVSRSHKTSRKLILNVNSNGDAVVEGQLETH 1112
 Db 1053 DGENNVPESSGRHGRGRTTKRHYRYSVSRSHKTSRKLILNVNSNGDAVVEGQLETH 1112
 Qy 895 NRKAVTFKFDGDNPEEIIATIMVNDPILAIIBESFVQVREIIEKADMLSEDEVSEY 954
 Db 895 NRKAVTFKFDGDNPEEIIATIMVNDPILAIIBESFVQVREIIEKADMLSEDEVSEY 954
 Qy 1113 NRKAVTFKFDGDNPEEIIATIMVNDPILAIIBESFVQVREIIEKADMLSEDEVSEY 1172
 Db 1113 NRKAVTFKFDGDNPEEIIATIMVNDPILAIIBESFVQVREIIEKADMLSEDEVSEY 1172
 Qy 955 EGDGLESLOQKDDYGSQKLGFEKQPIIPASSMPOQIGIPSSSLTQVHSAQRRTIV 1014
 Db 955 EGDGLESLOQKDDYGSQKLGFEKQPIIPASSMPOQIGIPSSSLTQVHSAQRRTIV 1014

Dh	1173	EGDGBEJLQCKDXDYGFSGQKLBEBFQOPIPBASMPQOIGIPPSISLTQVYVHSAARREIV	1232
QY	1015	SPVBSRLRESKVPFSEITDTVYAASTAOSPGMNLSHSASLSLQOAFSELRAAQWTEGPN	1074
Dh	1233	SPVPSRSLRESKVPFSEITDTVYAASTAOSPGMNLSHSASLSLQOAFSELRAAQWTEGPN	1232
QY	1075	TAPRPFHSHTGPFPVVPFPLSSIAQVPTTAAATAVPATSSPNDISTSVIOSEYVUTE	1134
Dh	1293	TAPRPFHSHTGPFPVVPFPLSSIAQVPTTAAATAVPATSSPNDISTSVIOSEYVUTE	1352
QY	1135	EGIAVAIASTGCVTSGGLPIPVSESPVLSSVSSITIPAVVISTTSPPSLQVPTSEI	1194
Dh	1353	EGIAVAIASTGCVTSGGLPIPVSESPVLSSVSSITIPAVVISTTSPPSLQVPTSEI	1412
QY	1195	VVSSITALPYVTASATSASAGGSTATPCKPDAVVVSQOAGSTTVGATILTSVSTTSPS	1254
Dh	1413	VVSSITALPYVTASATSASAGGSTATPCKPDAVVVSQOAGSTTVGATILTSVSTTSPS	1472
QY	1255	TASQSLTQSSSTSTPPLAETVYVVAHSLDTHSHSTTGCLAFSLSAPSSSSPPAGVSSY	1314
Dh	1473	TASQSLTQSSSTSTPPLAETVYVVAHSLDTHSHSTTGCLAFSLSAPSSSSPPAGVSSY	1532
QY	1315	ISOPEGHLPVLPVPSVIASTPILPQAGSTPPLPQVPSIPPLQPVANVAVOQTLIHS	1374
Dh	1533	ISOPEGHLPVLPVPSVIASTPILPQAGSTPPLPQVPSIPPLQPVANVAVOQTLIHS	1592
QY	1375	QPOPALLENQHTHCPEVSDPTOPKAPGIDDIKLEEKRLSFBSSSSGAQHASVLET	1434
Dh	1593	QPOPALLENQHTHCPEVSDPTOPKAPGIDDIKLEEKRLSFBSSSSGAQHASVLET	1652
QY	1435	SLVIESTVTPGIPPTTAAVAPSKULNSTSTCIPRPNLPIGYALVYTPVYTPGQVSTVST	1494
Dh	1653	SLVIESTVTPGIPPTTAAVAPSKULNSTSTCIPRPNLPIGYALVYTPVYTPGQVSTVST	1712
QY	1495	TTSQGVKPEPTAASKPPLTAKAPVLPVOTELPAGTLPSEOLPERPGSLTQSOQPLEDLAOL	1554
Dh	1713	TTSQGVKPEPTAASKPPLTAKAPVLPVOTELPAGTLPSEOLPERPGSLTQSOQPLEDLAOL	1772
QY	1555	RRTLSPEXITVTSAGVPVSMAPTAITBAGTOPQKGVSQVKEGVLATSSGAGVFKMGKF	1614
Dh	1773	RRTLSPEXITVTSAGVPVSMAPTAITBAGTOPQKGVSQVKEGVLATSSGAGVFKMGKF	1832
QY	1615	QVSTVAADGAQKEGKNSKSDAKSVHFESSTSSSVLSSSPBSTLYKBPNGITIPGISD	1674
Dh	1833	QVSTVAADGAQKEGKNSKSDAKSVHFESSTSSSVLSSSPBSTLYKBPNGITIPGISD	1892
QY	1675	VPSBAHKTTTAEAKSDPTGQPTKVGRFOVTTTANKVGRFVSUKTEBDKLTTPRKEGVPVSP	1734
Dh	1893	VPSBAHKTTTAEAKSDPTGQPTKVGRFOVTTTANKVGRFVSUKTEBDKLTTPRKEGVPVSP	1952
QY	1735	FMDLEQAVLPVAVIPKKEKPELSEPSHLNGPSSDPEAFLSRDVDGSGSPHSPIQLSSKS	1794
Dh	1953	FMDLEQAVLPVAVIPKKEKPELSEPSHLNGPSSDPEAFLSRDVDGSGSPHSPIQLSSKS	2012
QY	1795	LPSONLSQSLNSFNSSYMSDNESDIEDBDLKLRLRDKHLKEIODIOSROKHEIES	1854
Dh	2013	LPSONLSQSLNSFNSSYMSDNESDIEDBDLKLRLRDKHLKEIODIOSROKHEIES	2072
QY	1855	LYTLKGKVPVAVIIPPAAPLSGRRRRPTKSGKSSSSSSLGKNSPOLSGNLGSOQASVY	1914
Dh	2073	LYTLKGKVPVAVIIPPAAPLSGRRRRPTKSGKSSSSSSLGKNSPOLSGNLGSOQASVY	2132
QY	1915	LHPQOULHPNGNIPESGONOLQFLKSPSSDNLYSAFSTSGAISVSLAPGCGTSTN	1974
Dh	2133	LHPQOULHPNGNIPESGONOLQFLKSPSSDNLYSAFSTSGAISVSLAPGCGTSTN	2192
QY	1975	TVGATTVNSQAAQOAPPAWTSRRKGTFTDOLHKLVDNARDAAMNLSGRGSKGHMNYEGPG	2034
Dh	2193	TVGATTVNSQAAQOAPPAWTSRRKGTFTDOLHKLVDNARDAAMNLSGRGSKGHMNYEGPG	2252
QY	2035	MARFESAIPGOLCISMTSNLGGSAIISAASATSLGHFTKSNCPPOQOYFPATPPGAQWGT	2094
Dh	2253	MARFESAIPGOLCISMTSNLGGSAIISAASATSLGHFTKSNCPPOQOYFPATPPGAQWGT	2312

OY	2095	GGPAPQPLGQFOPVGTASLQNPNISNLOKSIINPQSGNLRTT	2136
Dd	2313	GGPAPQPLGQFOPVGTASLQNPNISNLOKSIINPQSGNLRTT	2354
		RESULT 10	
		ABO444403	
		ID ABO444403 standard; protein; 2076 AA.	
		XX	
		AC ABO444403;	
		XX	
		DT 26-SEP-2003 (first entry)	
		XX	
		DE Novel human protein kinase #23.	
		XX	
		KW Human; kinase; enzyme; cosmetic application; nutraceutical application.	
		OS Homo sapiens.	
		XX	
		PN US6541252-B1.	
		XX	
		PD 01-APR-2003.	
		XX	
		PF 14-MAY-2001; 2001US-00854856.	
		XX	
		PR 19-MAY-2000; 2000US-0206015P.	
		XX	
		(LEXI-) LEXICON GENETICS INC.	
		PA	
		PI Walke DW, Hilburn E, Donoho G, Turner CA;	
		XX	
		DR MPI; 2003-575927/54.	
		N-PSDB; ACH03786.	
		PT New nucleic acid encoding novel human proteins; useful in cosmetic and	
		XX	
		nutraceutical applications.	
		PS Disclosure; Page; 11pp; English.	
		XX	
		The invention relates to a new isolated nucleic acid encoding a novel	
		CC human protein kinase. The nucleic acid is useful in cosmetic and	
		CC nutritional applications. The present sequence represents the amino	
		CC acid sequence of a novel human protein kinase. Note: The sequence data	
		CC for this patent did not from part of the printed specification but was	
		CC obtained in electronic format directly from USPTO at	
		CC seqdata.uspto.gov/sequence.html?docID=6541252B1	
		XX	
		SQ Sequence 2076 AA;	
		Query Match 97.2%; Score 10508; DB 7; Length 2076;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 2076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	61	MDKDSRGAAATTTTTHRRFFRRSVICDNNATALELPGLPLSLPOPSIPAAVPQSAPPEPH	120
Dd	1	MDKDSRGAAATTTTTHRRFFRRSVICDNNATALELPGLPLSLPOPSIPAAVPQSAPPEPH	60
OY	121	REELVTATATSQVAQQCPAAAAPEEQAVAGPAPTSSSTSKORPVQSPISVGSKEPPP	180
Dd	61	REELVTATATSQVAQQCPAAAAPEEQAVAGPAPTSSSTSKORPVQSPISVGSKEPPP	120
OY	181	ARSGSGGSAKEPEERSQODDIIELETAVGMSNDGRFLKPDIETGRSGFTVVGGLD	240
Dd	121	ARSGSGGSAKEPEERSQODDIIELETAVGMSNDGRFLKPDIETGRSGFTVVGGLD	180
OY	241	TETVEVAMCELDQRKLTKSERQRFKEAEMLKQLQHPIIVRFYSWSESTVKGKKCIVLV	300
Dd	181	TETVEVAMCELDQRKLTKSERQRFKEAEMLKQLQHPIIVRFYSWSESTVKGKKCIVLV	240
OY	301	TEIMTSGLTKTYIKRFVVMKIIVLRSCROILKGLQELHTRTPPIIHRLDKCDNIPIITGP	360
Dd	241	TEIMTSGLTKTYIKRFVVMKIIVLRSCROILKGLQELHTRTPPIIHRLDKCDNIPIITGP	300

Qy	361	TGSVKIGDLGLATLKRASFPAKSVIGTPEFMAPEWEEKYDSVDVYAFGMCMLEMAISEY	420
Dd	301	TGSVKIGDLGLATLKRASFPAKSVIGTPEFMAPEWEEKYDSVDVYAFGMCMLEMAISEY	360
Qy	421	PYSECQNAAQIYRRVTSVGPASPDKVAIPVKEIIIEGCIQNKDERYSIKDLINHAFFQ	480
Dd	361	PYSECQNAAQIYRRVTSVGPASPDKVAIPVKEIIIEGCIQNKDERYSIKDLINHAFFQ	420
Qy	481	EETGVARELAEDDEGKAIKLMRLIEDIKLKGKYNONEAIEFSFDLERVPEPDVQEM	540
Dd	421	EETGVARELAEDDEGKAIKLMRLIEDIKLKGKYNONEAIEFSFDLERVPEPDVQEM	480
Qy	541	VESGVYEGDHKTAKAKIKRVSILIKRREORQLVREOEKKKEESLAKQVQSSASQ	600
Dd	481	VESGVYEGDHKTAKAKIKRVSILIKRREORQLVREOEKKKEESLAKQVQSSASQ	540
Qy	601	TGIRQLPSASTGITPTASTASVSTQVPEPEADQHQQLQYQPSISVLSGTVDSGQ	660
Dd	541	TGIRQLPSASTGITPTASTASVSTQVPEPEADQHQQLQYQPSISVLSGTVDSGQ	600
Qy	661	SSVFTESVSSQQTAVSGSOHEOAHSTGVPGHIPSITQAOQOPHGVPPSSVAGQSGQ	720
Dd	601	SSVFTESVSSQQTAVSGSOHEOAHSTGVPGHIPSITQAOQOPHGVPPSSVAGQSGQ	660
Qy	721	QPSSSLTGVSSSQPIQHPQQQOGIQOTAPPOQTVOYSLSQTSSEATTAQPPSOQAP	780
Dd	661	QPSSSLTGVSSSQPIQHPQQQOGIQOTAPPOQTVOYSLSQTSSEATTAQPPSOQAP	720
Qy	781	QVLPOVSAKOSTQGVSCVAPAEVAVAAQOPATOPTILLASVDSAHSDVAGSGDGENV	840
Dd	721	QVLPOVSAKOSTQGVSCVAPAEVAVAAQOPATOPTILLASVDSAHSDVAGSGDGENV	780
Qy	841	PSSSGRHEGRRTTKHYKRSVSRHEKTSRPKRLINVMKGRVYECQLETNRKMYT	900
Dd	781	PSSSGRHEGRRTTKHYKRSVSRHEKTSRPKRLINVMKGRVYECQLETNRKMYT	840
Qy	901	FKPDLIDGNPEELATIMWNNDFILAIERESFVDQVREIIEKADMLSEDVSVEEBQGL	960
Dd	841	FKPDLIDGNPEELATIMWNNDFILAIERESFVDQVREIIEKADMLSEDVSVEEBQGL	900
Qy	961	ESLQKDDYFGSSQKLEGEFKQPIPASSMPQOIGIPTSSLTQVYVHSAGRFFVSPVES	1020
Dd	901	ESLQKDDYFGSSQKLEGEFKQPIPASSMPQOIGIPTSSLTQVYVHSAGRFFVSPVES	960
Qy	1021	RLRSKVPFSEITDTVAASTASQPMNLSSHASSLSLQOAFSELRRAQMTGEPNTAPNF	1080
Dd	961	RLRSKVPFSEITDTVAASTASQPMNLSSHASSLSLQOAFSELRRAQMTGEPNTAPNF	1020
Qy	1081	SHGTPTFPVVPPLSLIAGVPTTAATAPVATSSPPNDISTSVIOSEVTVPTEEGIAGV	1140
Dd	1021	SHGTPTFPVVPPLSLIAGVPTTAATAPVATSSPPNDISTSVIOSEVTVPTEEGIAGV	1080
Qy	1141	ATSTGVVTSGLPLIPVSESPVLSVSSITIPAVVSIITSSPSLQVPTSEIIVSSTA	1200
Dd	1081	ATSTGVVTSGLPLIPVSESPVLSVSSITIPAVVSIITSSPSLQVPTSEIIVSSTA	1140
Qy	1201	LYPSVTASATASAGSATPGRKPPAVVSOQAAGSTTVGATLTVSVSTTSSFPETAOLS	1260
Dd	1141	LYPSVTASATASAGSATPGRKPPAVVSOQAAGSTTVGATLTVSVSTTSSFPETAOLS	1200
Qy	1261	IQSSSTSTPTLATVTVVSAHSLDKTSHSSTTGLAFSLASAPSSSSPQAGVSVYSIOPG	1320
Dd	1201	IQSSSTSTPTLATVTVVSAHSLDKTSHSSTTGLAFSLASAPSSSSPQAGVSVYSIOPG	1260
Qy	1321	LHPVIVSVIATSTPLPQAGPTSTPLIPQVPSIPPLVQVAVNPVAVQOQTLHISQOPAL	1380
Dd	1261	LHPVIVSVIATSTPLPQAGPTSTPLIPQVPSIPPLVQVAVNPVAVQOQTLHISQOPAL	1320
Qy	1381	LPNQPHHCPVDDTQPKAGIDDKLEKLSLSESHSSSGAQAHSVLSFVSVES	1440
Dd	1321	LPNQPHHCPVDDTQPKAGIDDKLEKLSLSESHSSSGAQAHSVLSFVSVES	1380

Qy	1441	TVTPIETTAVAPSKLITSTTSTCLPPTNLPLGVALPVTVPVQGVSTPVSITTSQV	1500
Dd	1381	TVTPIETTAVAPSKLITSTTSTCLPPTNLPLGVALPVTVPVQGVSTPVSITTSQV	1440
Qy	1501	PGTAPSKPLTKAVLPVGTBLPAGTLPSEQLPFPFGPSLTSQOQPLEDLDAQLRRTLS	1560
Dd	1441	PGTAPSKPLTKAVLPVGTBLPAGTLPSEQLPFPFGPSLTSQOQPLEDLDAQLRRTLS	1500
Qy	1561	EXTIVTSVAVPVSAAATLATEGTQPOKQVSVQKEBPVLTATSSGACVFKMGRQVSVAA	1620
Dd	1501	EXTIVTSVAVPVSAAATLATEGTQPOKQVSVQKEBPVLTATSSGACVFKMGRQVSVAA	1560
Qy	1621	DGAQKEGKNKSEDAKSHFESSSTESSVLSSSPSESTLVKPEPGITIPGISDVPESAH	1680
Dd	1561	DGAQKEGKNKSEDAKSHFESSSTESSVLSSSPSESTLVKPEPGITIPGISDVPESAH	1620
Qy	1681	KTINSEAKSDTQGTQVGRFOVTTANKVGRFSVSKTEBKITDTYKKEGPVAPPPMDLEQ	1740
Dd	1621	KTINSEAKSDTQGTQVGRFOVTTANKVGRFSVSKTEBKITDTYKKEGPVAPPPMDLEQ	1680
Qy	1741	AVLPAVLPKKEKPELSPSHLNGPSSDPEAAFLSRDVGSGSPHOLSSKSLPSQNL	1800
Dd	1681	AVLPAVLPKKEKPELSPSHLNGPSSDPEAAFLSRDVGSGSPHOLSSKSLPSQNL	1740
Qy	1801	SQSLSNSFNSSVMSDNESEDIEDEDLKLELRRLRDKMLKEIQDQSRQKHEIESLYTKLG	1860
Dd	1741	SQSLSNSFNSSVMSDNESEDIEDEDLKLELRRLRDKMLKEIQDQSRQKHEIESLYTKLG	1800
Qy	1861	KVPPAVIIPPAAPLSGRRRPPTYSKGSKSSRSSSLGNKSPQLSGNLSGQSAASYLHPQOT	1920
Dd	1801	KVPPAVIIPPAAPLSGRRRPPTYSKGSKSSRSSSLGNKSPQLSGNLSGQSAASYLHPQOT	1860
Qy	1921	LHPGNTPESSGONOLQPLKPSPSDNLVYAFISDGAISVPSISAPQOGTSTTVGATV	1980
Dd	1861	LHPGNTPESSGONOLQPLKPSPSDNLVYAFISDGAISVPSISAPQOGTSTTVGATV	1920
Qy	1981	NSQAQAQAPPAAMTSSRKGTFTDDLAKLVDMWARDAMNLSGRGSGKGMNYEGPMARKFS	2040
Dd	1921	NSQAQAQAPPAAMTSSRKGTFTDDLAKLVDMWARDAMNLSGRGSGKGMNYEGPMARKFS	1980
Qy	2041	APGQLCISMTSNUGSAPISAASATSLGHTTKSCPCPQOYGFPAITPPGAOMSGTGAPAPQ	2100
Dd	1981	APGQLCISMTSNUGSAPISAASATSLGHTTKSCPCPQOYGFPAITPPGAOMSGTGAPAPQ	2040
Qy	2101	PLGQFQPVGTASLQENENISNLQKSISSNPQSNLRTT	2136
Dd	2041	PLGQFQPVGTASLQENENISNLQKSISSNPQSNLRTT	2101

RESULT 11			
XX	ABO44400	ID	ABO44400 standard; protein: 2169 AA.
XX	ABO44400;	AC	
XX	26-SEP-2003	DT	(first entry)
XX	Novel human protein kinase #20.	DE	
XX	Human; kinase; enzyme; cosmetic application; nutraceutical application.	KW	
XX	Homo sapiens.	OS	
XX	US6541252-B1.	PN	
XX	01-APR-2003.	PD	
XX	14-MAY-2001; 2001US-00854856.	PF	
XX	19-MAY-2000; 2000US-0206015P.	PR	
XX	(TEXT-) LEXICON GENNETTICS INC.	PA	

PI Walke DW, Hilbun E, Donoho G, Turner CA;
 XX MPI: 2003-575927/54.
 DR N-PSDB; ACH03783.
 XX
 XX New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutritional applications.
 PT
 XX
 PS Disclosure; Page; 11pp; English.
 XX
 XX The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutritional applications. The present sequence represents the amino
 CC acid sequence of a novel human protein kinase. Note: The sequence data
 CC for this patent did not form part of the printed specification but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=6541252B1
 CC
 XX
 SQ Sequence 2169 AA:

Query Match 96.7%; Score 10451.5; DB 7; Length 2169;
 Best local similarity 95.7%; Pred. No. 0;
 Matches 2076; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

QY 61 MDXSRGAATTTTTEHFRFRSVICDSNATALELPLSLPQPSIPAAVPOSAPPPEH 120
 Db 1 MDXSRGAATTTTTEHFRFRSVICDSNATALELPLSLPQPSIPAAVPOSAPPPEH 60
 QY 121 REETVATATISQVAQPPAAAPGEOAVAGPAPSTVPSTSKDRPVSGPSLVGSKKEEPP 180
 Db 61 REETVATATISQVAQPPAAAPGEOAVAGPAPSTVPSTSKDRPVSGPSLVGSKKEEPP 120
 QY 181 ARSGSGGSAKEPEEESQODDIELEETKAVGMSNDREFLKPFIETGRGSEFKTVYGLD 240
 Db 121 ARSGSGGSAKEPEEESQODDIELEETKAVGMSNDREFLKPFIETGRGSEFKTVYGLD 180
 QY 241 TETVEVAMCELQDRKLTSEKORFKEEAEMIKGIQHNIVRFYDSMESTVKGKCIYLV 300
 Db 181 TETVEVAMCELQDRKLTSEKORFKEEAEMIKGIQHNIVRFYDSMESTVKGKCIYLV 240
 QY 301 TELMTSGTLKTYLKRFKVMKIKVLRSCROILKGIQFHTPTPIIHRDLKCDNIPTTGP 360
 Db 241 TELMTSGTLKTYLKRFKVMKIKVLRSCROILKGIQFHTPTPIIHRDLKCDNIPTTGP 300
 QY 361 TGSVKIGLGLATIKRASFASVIGTPEFMAPEMVEEKYDESVDVYAFGMCMLMATSEY 420
 Db 301 TGSVKIGLGLATIKRASFASVIGTPEFMAPEMVEEKYDESVDVYAFGMCMLMATSEY 360
 QY 421 PYSECQNAOQYRRVTSGVFPKVAIPVKEKIEGCIKQNDERSIKDLNHAFFQ 480
 Db 361 PYSECQNAOQYRRVTSGVFPKVAIPVKEKIEGCIKQNDERSIKDLNHAFFQ 420
 QY 481 BETGVAVELAEEDGEEKIAIKLMLRIEDIKLKGGYKQNEAIEFSFDLERDVPEDVAQEM 540
 Db 421 BETGVAVELAEEDGEEKIAIKLMLRIEDIKLKGGYKQNEAIEFSFDLERDVPEDVAQEM 480
 QY 541 VESGVYCEGDHMKAKAIKQVSLIKRRBORQVRESEKKKQNEESLTKQVSSASQ 600
 Db 481 VESGVYCEGDHMKAKAIKQVSLIKRRBORQVRESEKKKQNEESLTKQVSSASQ 540
 QY 601 TGIKOLPSASTGISTASVSTOVEPEPEADQHQLOYOQPSISVLSDGVDSGQG 660
 Db 541 TGIKOLPSASTGISTASVSTOVEPEPEADQHQLOYOQPSISVLSDGVDSGQG 600
 QY 661 SSVTETESRVSQQTIVSYSGQHEQAHSTGVGHIPSTVYQAOQPHGYPPSSVAQSGSG 720
 Db 601 SSVTETESRVSQQTIVSYSGQHEQAHSTGVGHIPSTVYQAOQPHGYPPSSVAQSGSG 660
 QY 721 QPSSSSTLVGSSSQPIQHPQOQOGIIOOTAPPOQTVOYSLQTSSTRTTQAPVSOQAP 780
 Db 661 QPSSSSTLVGSSSQPIQHPQOQOGIIOOTAPPOQTVOYSLQTSSTRTTQAPVSOQAP 720
 QY 781 QVLEFQVSAKQO----- 791

Db 721 QVLEFQVSAKQOFPPLRPPQYPCDSNATAPSSNVAVCIHSTVLXPMPETEVLATPGYEPT 780
 QY 792 -----STQVSGVAPAEPAV 807
 Db 781 VVQPYVESNLLVPMGVGGVQVQVSGGSLAQAPLTSSQQAFLVSTQGVSGVAPAEPAV 840
 QY 808 AQPOATQPTTLASVDNAHSDVAGMDGNENNVSSEGRHGGRTTKHRYKSVSRSHHE 867
 Db 841 AQPOATQPTTLASVDNAHSDVAGMDGNENNVSSEGRHGGRTTKHRYKSVSRSHHE 800
 QY 868 KTSAPKRIINLVNSKGRVVECOLETHNRKMTVEFKFDGNGPEIATIMNNPILIE 927
 Db 901 KTSAPKRIINLVNSKGRVVECOLETHNRKMTVEFKFDGNGPEIATIMNNPILIE 860
 QY 928 RESFVDQREIIEKADMLSEDVSEPEGDGLSLQKDDYGFSGSKLEGEFKQPIPA 987
 Db 961 RESFVDQREIIEKADMLSEDVSEPEGDGLSLQKDDYGFSGSKLEGEFKQPIPA 920
 QY 988 SSMPOQIGIPSSLTQVHSAGRFIVSPVESRLRESKVPSEITDVAASTAQSPGMN 1047
 Db 1021 SSMPOQIGIPSSLTQVHSAGRFIVSPVESRLRESKVPSEITDVAASTAQSPGMN 980
 QY 1048 LSHSASSLSIQQAASELRRAQMTGEPNTAPENFSGHTGTFPVVPPFLSSIAGVPTTAAT 1107
 Db 1081 LSHSASSLSIQQAASELRRAQMTGEPNTAPENFSGHTGTFPVVPPFLSSIAGVPTTAAT 1040
 QY 1108 APVPATSSPNDISTSVIOSEVIVTEBGIAGVATSGVTSQGLPIPVSESVLSSV 1167
 Db 1141 APVPATSSPNDISTSVIOSEVIVTEBGIAGVATSGVTSQGLPIPVSESVLSSV 1100
 QY 1168 SSTIPAVASISTSPLOVPTSEIVSSSTALYPSTVATASAGSGSTATGPKRPA 1227
 Db 1201 SSTIPAVASISTSPLOVPTSEIVSSSTALYPSTVATASAGSGSTATGPKRPA 1160
 QY 1228 VVSQQAAGSTTVGATLTVSVSTTTSFPSTASQSLTQSSSTSTPTLAETVVAHSLDKTS 1287
 Db 1261 VVSQQAAGSTTVGATLTVSVSTTTSFPSTASQSLTQSSSTSTPTLAETVVAHSLDKTS 1220
 QY 1288 HSSTTGIAFCSAPSSSSSPGAGVSYISQPGGHPVIVBVIASSTPLPAAAGTSTPL 1347
 Db 1321 HSSTTGIAFCSAPSSSSSPGAGVSYISQPGGHPVIVBVIASSTPLPAAAGTSTPL 1280
 QY 1348 LPQVPSIPPLVOPAVANPAVOQTLIHQOPALPNOPTHCPVDSTQKAGIDIDIK 1407
 Db 1381 LPQVPSIPPLVOPAVANPAVOQTLIHQOPALPNOPTHCPVDSTQKAGIDIDIK 1340
 QY 1408 TLEBKLSLSEHSSSGAASHVSLSTSVLESTVTEGIPPTTAAVAPSKLITSTSTCLP 1467
 Db 1441 TLEBKLSLSEHSSSGAASHVSLSTSVLESTVTEGIPPTTAAVAPSKLITSTSTCLP 1400
 QY 1468 TNLPLGTVALPVTPVTPGQVSTVSTTTGKVGKGTAPSKPPLKAPVLPYGTLPAGTL 1527
 Db 1501 TNLPLGTVALPVTPVTPGQVSTVSTTTGKVGKGTAPSKPPLKAPVLPYGTLPAGTL 1460
 QY 1528 PSBQLPFPFGPSLQSQOPLDADQRLTSLSPXITVTSVGVGVMAAPATIEAGTQF 1587
 Db 1561 PSBQLPFPFGPSLQSQOPLDADQRLTSLSPXITVTSVGVGVMAAPATIEAGTQF 1520
 QY 1588 QKGVSYQKEGVLAATSSGAGVFKMGRFQVVAADQAKGKNGKSDAKSVHFSSTSESS 1647
 Db 1621 QKGVSYQKEGVLAATSSGAGVFKMGRFQVVAADQAKGKNGKSDAKSVHFSSTSESS 1580
 QY 1648 VLSASSPESTLVKRENGITIPGISVVPESBAKTTTSEAKSDPTKGRFQVTTAN 1707
 Db 1681 VLSASSPESTLVKRENGITIPGISVVPESBAKTTTSEAKSDPTKGRFQVTTAN 1640
 QY 1708 KVGKFSYKTEDEKTTDPKKGPPVAPPFMDLEQAVLPAVLPKKEKPELSESHLNGSSSD 1767
 Db 1741 KVGKFSYKTEDEKTTDPKKGPPVAPPFMDLEQAVLPAVLPKKEKPELSESHLNGSSSD 1700
 QY 1768 PEAFLSRDVEDGSGSPHSPHOLSKSLPQONTLSQSLNSFNSSYMSDNESDIIEDEDLK 1827

Db 1801 PEAALFSDVDGSGSPHSPOLSKSLPSQNLSSQSLNSFNSSYSSSDNEDEBDLK 1860
 QY 1828 LEIARLRDPKHLEIIDLQSRQKHIEISLYTKLGKVPVAVIIPPAAPISGRRRPTKSGS 1887
 Db 1861 LEIARLRDPKHLEIIDLQSRQKHIEISLYTKLGKVPVAVIIPPAAPISGRRRPTKSGS 1920
 QY 1888 KSSRSSSLGNKSPOLSGNLSCQSAASVLAHPQOTLHPNIPESGONLLQELKSPSSDN 1947
 Db 1921 KSSRSSSLGNKSPOLSGNLSCQSAASVLAHPQOTLHPNIPESGONLLQELKSPSSDN 1980
 QY 1948 LYSAFSTSGAISVSLSLAPGQGTSTNTVGATVNSQAQAOPPMATSSRKGTFTDDLHL 2007
 Db 1981 LYSAFSTSGAISVSLSLAPGQGTSTNTVGATVNSQAQAOPPMATSSRKGTFTDDLHL 2040
 QY 2008 VDNARADAMNLGSRGSGKHNNVYEGPGWARKFSAAGOLCISMTSNLGGADISAASATSL 2067
 Db 2041 VDNARADAMNLGSRGSGKHNNVYEGPGWARKFSAAGOLCISMTSNLGGADISAASATSL 2100
 QY 2068 GHFTKSMCPPOQYGFPAIPFGAOWSGTGPAPOPPLQGPQVPVGTASLQNFNISLQKSISSN 2127
 Db 2101 GHFTKSMCPPOQYGFPAIPFGAOWSGTGPAPOPPLQGPQVPVGTASLQNFNISLQKSISSN 2160
 QY 2128 PPGSNLRTT 2136
 Db 2161 PPGSNLRTT 2169

RESULT 12

ABO44397 standard; protein; 2322 AA.

ID ABO44397 standard; protein; 2322 AA.
 AC ABO44397;
 XX 26-SEP-2003 (first entry)
 DT Novel human protein kinase #17.
 DE Human; kinase; enzyme; cosmetic application; nutraceutical application.
 KW Homo sapiens.
 OS US6541252-B1.
 EN 01-APR-2003.
 PD 14-MAY-2001; 2001US-00854856.
 PF 19-MAY-2000; 2000US-0206015P.
 ER (LEXI-) LEXICON GENETICS INC.
 XX PA
 PI Walke DW, Hilbun E, Donoho G, Turner CA;
 DR WPI; 2003-575927/54.
 DR N-PSDB; ACH03780.
 XX
 PT New nucleic acid encoding novel human proteins, useful in cosmetic and
 PT nutraceutical applications.
 PS Disclosure; Page; 11pp; English.
 CC The invention relates to a new isolated nucleic acid encoding a novel
 CC human protein kinase. The nucleic acid is useful in cosmetic and
 CC nutraceutical applications. The present sequence represents the amino
 CC acid sequence of a novel human protein kinase. Note: The sequence data
 CC for this patent did not form part of the printed specification but was
 CC obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=6541252B1
 XX Sequence 2322 AA:

Query Match 96.0%; Score 10375; DB 7; Length 2322;
 Best Local Similarity 89.4%; Pred. No. 0;

	Matches 2076;	Conservative 0;	Mismatches 0;	Indels 246;	Gaps 1;
QY 61	MDKSRGAAATTTTTERFRFRSVICDNNATLELPGIPLSLPOPSTIPAAVPOGAPPEPH				120
Db 1	MDKSRGAAATTTTTERFRFRSVICDNNATLELPGIPLSLPOPSTIPAAVPOGAPPEPH				60
QY 121	REETVATATSOVAQAPAAAAPGEQAVAGAPSTVSTSKDBPVSOPLVSGKEEPP				180
Db 61	REETVATATSOVAQAPAAAAPGEQAVAGAPSTVSTSKDBPVSOPLVSGKEEPP				120
QY 181	ARSSGGGSAKEPEEESQOQDDIELEETAVAGSNNGRFLKEPIELIGRGSFKTVYKLD				240
Db 121	ARSSGGGSAKEPEEESQOQDDIELEETAVAGSNNGRFLKEPIELIGRGSFKTVYKLD				180
QY 241	TEETVEAMCELOQRKLTSPROFKEAEMLKGLQHPNIVRFSDMESYTKGKCIYLV				300
Db 181	TEETVEAMCELOQRKLTSPROFKEAEMLKGLQHPNIVRFSDMESYTKGKCIYLV				240
QY 301	TEMTSGTLKTYLKRFRVMKIKVLRSWCROILKGLQELHTPTPIIHRDLKCDNIFITGP				360
Db 241	TEMTSGTLKTYLKRFRVMKIKVLRSWCROILKGLQELHTPTPIIHRDLKCDNIFITGP				300
QY 361	TGSVKIGDLGLATIKRASPAKSVIGTEPFMAPEMYEKEKYDESUVVYAFGMCMLEMAISEY				420
Db 301	TGSVKIGDLGLATIKRASPAKSVIGTEPFMAPEMYEKEKYDESUVVYAFGMCMLEMAISEY				360
QY 421	PYSECQNAAOIYRRTSGVXPASFEKVAIPPEKEIEIGCTIRKNDERSYIDLNLHAFQ				480
Db 361	PYSECQNAAOIYRRTSGVXPASFEKVAIPPEKEIEIGCTIRKNDERSYIDLNLHAFQ				420
QY 481	EETGVRVLAEDDGEKIAIKLMLRIDIKLKQKYKNDENAIIESFDLERDVEDVAQEM				540
Db 421	EETGVRVLAEDDGEKIAIKLMLRIDIKLKQKYKNDENAIIESFDLERDVEDVAQEM				480
QY 541	VEGTVGCEGDKHTAKAKIKDRVSLIKRREKROLVREBEKKKEESSLKQOVQSSAQ				600
Db 481	VEGTVGCEGDKHTAKAKIKDRVSLIKRREKROLVREBEKKKEESSLKQOVQSSAQ				540
QY 601	TGIRQLPSASTGTIPTASTSASVSTQVEPEPEPADQHQLOLOQOPSTISVLSGTVDSQ				660
Db 541	TGIRQLPSASTGTIPTASTSASVSTQVEPEPEPADQHQLOLOQOPSTISVLSGTVDSQ				600
QY 661	SSVFTESRVSSQOTVSYGSGHEQASHSTGTVPGHIPSTVQASQPHGVYPPSSVAQOGSQ				720
Db 601	SSVFTESRVSSQOTVSYGSGHEQASHSTGTVPGHIPSTVQASQPHGVYPPSSVAQOGSQ				660
QY 721	QPSSSILTVSSSQPIQHPOQOQCIQOTAPQOQVQYSLQSTSSSEATTNQPVSQOPAP				780
Db 661	QPSSSILTVSSSQPIQHPOQOQCIQOTAPQOQVQYSLQSTSSSEATTNQPVSQOPAP				720
QY 781	QVLPOVSAGKQ-----				791
Db 721	QVLPOVSAGKQLPVSPVPTTQGEPOIPVATQPSVVPVHSGAHLPVGQPLPTLLPQY				780
QY 792	-----				791
Db 781	VSQPISTPHVSTAQTGSSLPITMAAGITQPLITLASSATTAAIPGVSTVPSQLPTLL				840
QY 792	-----				791
Db 841	QPTVLPSQVHPOLLQPAVQSGMGI PANLQAAEVLSSGDLVYGFPPRLPQYRGDSNI				900
QY 792	-----				791
Db 901	APSSNVASVCIHSTVLAHPMTVEIATPGVPTVQVPEVSNLIVPMGQVGGQVQVSPG				960
QY 792	-----				834
Db 961	GSIAQAPTTSSQAVLESTQGVSVQVAPAEVAVAVQPAQTQPTLLASSVDSASHSVASGMS				1020
QY 835	DCNENVPSSGRRHGRTTKRHVRSKVSRSRHEKTSRPKRIILVNSNKGDPVVECOLETH				894
Db 1021	DGNENVPSSGRRHGRTTKRHVRSKVSRSRHEKTSRPKRIILVNSNKGDPVVECOLETH				1080


```

QY 895 NRKWTFFKFDLDGNDPERIATIMVNDPILAIERESFVDQVRELLIEKADENLSEVSEP 954
Db 1081 NRKWTFFKFDLDGNDPERIATIMVNDPILAIERESFVDQVRELLIEKADENLSEVSEP 1140
QY 955 EGDQGLJESLQGDNDYGFSGSQCLBEGEFKQPIPASMPQOIGIPISLTQVHSAGRFIV 1014
Db 1141 EGDQGLJESLQGDNDYGFSGSQCLBEGEFKQPIPASMPQOIGIPISLTQVHSAGRFIV 1200
QY 1015 SPVPSRRLRESKVFSEITDITVAASTAOSPGMNLSSHASLSLQCAFSELRRAQMTBGN 1074
Db 1201 SPVPSRRLRESKVFSEITDITVAASTAOSPGMNLSSHASLSLQCAFSELRRAQMTBGN 1260
QY 1075 TAPPEFSTGTFEPVPPFELSIAGVPTTAATAATVPATSSPNDISTSVIOSEVTEPTE 1134
Db 1261 TAPPEFSTGTFEPVPPFELSIAGVPTTAATAATVPATSSPNDISTSVIOSEVTEPTE 1320
QY 1135 EGIAGVATSTGVTSGGLPIPPVSESPLYSSVSSITTPAVSISTTSPLQVPTSTSEI 1194
Db 1321 EGIAGVATSTGVTSGGLPIPPVSESPLYSSVSSITTPAVSISTTSPLQVPTSTSEI 1380
QY 1195 VVSTALYPSVTVAATSAAGSTATPGKPPAVVSOQAAGSTTVGATLTSTSTTSPS 1254
Db 1381 VVSTALYPSVTVAATSAAGSTATPGKPPAVVSOQAAGSTTVGATLTSTSTTSPS 1440
QY 1255 TASQLSIQLSSSTSTPTLAETVWVAHSLDKTSHSTTGIAFSLAPSSSSPGAGVSSY 1314
Db 1441 TASQLSIQLSSSTSTPTLAETVWVAHSLDKTSHSTTGIAFSLAPSSSSPGAGVSSY 1500
QY 1315 ISQPGHPLVIVPSVIASTPILPQAAGPTSTPLPQVPSIPPLVQPVANVAVOQTLIHS 1374
Db 1501 ISQPGHPLVIVPSVIASTPILPQAAGPTSTPLPQVPSIPPLVQPVANVAVOQTLIHS 1560
QY 1375 QOPRALPLNQHHTHCPEDVSDTOPKABGIDIKITEELKRLSIFSHSSSGAQHSAVSLET 1434
Db 1561 QOPRALPLNQHHTHCPEDVSDTOPKABGIDIKITEELKRLSIFSHSSSGAQHSAVSLET 1620
QY 1435 SLVIESTYTPGIPITTAVAVSKLITSTSTCLPPTNLPGTVALPPTPVTEGGVSTPST 1494
Db 1621 SLVIESTYTPGIPITTAVAVSKLITSTSTCLPPTNLPGTVALPPTPVTEGGVSTPST 1680
QY 1495 TTSGVKEGTABSKPPLTKAPVLPVGTTELPACTLSEQLPFPFGSLTQSQPLEDLDAQL 1554
Db 1681 TTSGVKEGTABSKPPLTKAPVLPVGTTELPACTLSEQLPFPFGSLTQSQPLEDLDAQL 1740
QY 1555 RRTISPEKITTSVAVGPVMAAPTAITEAGTQPKQVSGVKEGVLATSSGAGVFKMGFR 1614
Db 1741 RRTISPEKITTSVAVGPVMAAPTAITEAGTQPKQVSGVKEGVLATSSGAGVFKMGFR 1800
QY 1615 QVSVAAADGAQKEGKNKSEDAKSVHFESTSESSVLSSSPSTLVKPEPNCITTPGISD 1674
Db 1801 QVSVAAADGAQKEGKNKSEDAKSVHFESTSESSVLSSSPSTLVKPEPNCITTPGISD 1860
QY 1675 VPESAHKTTASEAKSDTGQPTKVGPFQYTTANKVGRFSVSKTEDKITDTKKEGPVAP 1734
Db 1861 VPESAHKTTASEAKSDTGQPTKVGPFQYTTANKVGRFSVSKTEDKITDTKKEGPVAP 1920
QY 1735 FMDLEQAVLPVITPKKEPELSEPSHLNGPSSDEPAFLSDVDVDSGSSPHSPQLSSKS 1794
Db 1921 FMDLEQAVLPVITPKKEPELSEPSHLNGPSSDEPAFLSDVDVDSGSSPHSPQLSSKS 1980
QY 1795 LPSQNLQSLSNSFNSSYMSSDNESDIEDDLKLELRRLRPDKHLKEIIDLQSRKHETES 1854
Db 1981 LPSQNLQSLSNSFNSSYMSSDNESDIEDDLKLELRRLRPDKHLKEIIDLQSRKHETES 2040
QY 1855 LYTLGKVPVAVIIIPAAPLSGRRRPTKSGKSSRSSSLGNKSPQLSGNLSCQASAV 1914
Db 2041 LYTLGKVPVAVIIIPAAPLSGRRRPTKSGKSSRSSSLGNKSPQLSGNLSCQASAV 2100
QY 1915 LHPQOULHPGNTPESSQNLQPLKSPSSDNTYSAFTSDGALSVPSLSAPCGTSTSTN 1974
Db 2101 LHPQOULHPGNTPESSQNLQPLKSPSSDNTYSAFTSDGALSVPSLSAPCGTSTSTN 2160

```

```

QY 1975 TVGATVNSQAAQAPPMATSSRKGTFTDDLHKLVDNMARDAMNLSGRGSKGMNYEGPG 2034
Db 2161 TVGATVNSQAAQAPPMATSSRKGTFTDDLHKLVDNMARDAMNLSGRGSKGMNYEGPG 2220
QY 2035 MARKFSAPGOLCISMTSNLIGSAPISASATSLGHFTSMCPPOQYGPATPFGAOMSGT 2094
Db 2221 MARKFSAPGOLCISMTSNLIGSAPISASATSLGHFTSMCPPOQYGPATPFGAOMSGT 2280
QY 2095 GGPAPOPLGCPQPVGTASLQNFNLSNLQKSLSNPPGSMLRTT 2136
Db 2281 GGPAPOPLGCPQPVGTASLQNFNLSNLQKSLSNPPGSMLRTT 2322

RESULT 13
ABO44411
ID ABO44411 standard; protein, 2048 AA.
XX
XX ABO44411;
AC
XX
XX 26-SEP-2003 (first entry)
DT
XX
XX Novel human protein kinase #31.
DE
XX
XX Human; kinase; enzyme; cosmetic application; nutraceutical application.
KW
XX
XX Homo sapiens.
OS
XX
XX US6541252-B1.
PN
XX
XX 01-APR-2003.
PD
XX
XX 14-MAY-2001; 2001US-00854856.
PP
XX
XX 19-MAY-2000; 2000US-0206015P.
PR
XX
XX (LEXI-) LEXICON GENETICS INC.
PA
XX
XX
XX
XX
XX Walke DW, Hilbun E, Donoho G, Turner CA;
PI
XX
XX WPI; 2003-575927/54.
DR
XX
XX N-PSDB; ACH03794.
DX
XX
XX
XX
XX New nucleic acid encoding novel human proteins, useful in cosmetic and
PT nutraceutical applications.
PS
XX
XX Disclosure; Page: 11pp; English.
PS
XX
XX
XX
XX The invention relates to a new isolated nucleic acid encoding a novel
CC human protein kinase. The nucleic acid is useful in cosmetic and
CC nutraceutical applications. The present sequence represents the amino
CC acid sequence of a novel human protein kinase. Note: The sequence data
CC for this patent did not from part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=6541252B1
CC
XX
XX
SQ
Sequence 2048 AA:

Query Match 95.7%; Score 10344; DB 7; Length 2048;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 2048; Conservative 0; Mismatches 0; Indels 28; Gaps 1;

```

QY 241 TETTVAVACELQDRUKLTSEKQRFKEEAEMLKGLQHPNIVRFYDSWESTVKGKCIIVL 300
 DB 181 TETTVAVACELQDRUKLTSEKQRFKEEAEMLKGLQHPNIVRFYDSWESTVKGKCIIVL 240
 QY 301 TELMTSGTLKTYLKGPKYMKIKVLRSMCRQILKGLQFHTHTPPRIHRDLKCDNIPTTGP 360
 DB 241 TELMTSGTLKTYLKGPKYMKIKVLRSMCRQILKGLQFHTHTPPRIHRDLKCDNIPTTGP 300
 QY 361 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPBMYEKEKYDESUVVYAFGMCLMATSEY 420
 DB 301 TGSVKIGDLGLATLKRAFPKSVIGTPEFMAPBMYEKEKYDESUVVYAFGMCLMATSEY 360
 QY 421 PYSCONAAQIYRVTSGVCPKSPASFDKVAIPEVKIEIGCIRQNDERYSIKDLINHAFFQ 480
 DB 361 PYSCONAAQIYRVTSGVCPKSPASFDKVAIPEVKIEIGCIRQNDERYSIKDLINHAFFQ 420
 QY 481 EETGVRAVELAEDDGEKIAIKLWLRIEDIKLKGKYNKNEAIEFSFDLERDVPEDVAQEM 540
 DB 421 EETGVRAVELAEDDGEKIAIKLWLRIEDIKLKGKYNKNEAIEFSFDLERDVPEDVAQEM 480
 QY 541 VESGVCEGDKHTMAKAIKDRVSLIKRREOROLVREBQKKQKQESSLKQVQSSASQ 600
 DB 481 VESGVCEGDKHTMAKAIKDRVSLIKRREOROLVREBQKKQKQESSLKQVQSSASQ 540
 QY 601 TGIKQLPASASTGIPASTTSASVSTOVEPEEPADQHQOIQOQPSISVSDGTVDGQGG 660
 DB 541 TGIKQLPASASTGIPASTTSASVSTOVEPEEPADQHQOIQOQPSISVSDGTVDGQGG 600
 QY 661 SSVTESTRVSSQOIVSYSGSQHEQAHSTGTVGHIPSTYQAOSSQHGIVPESVVAQGSQGG 720
 DB 601 SSVTESTRVSSQOIVSYSGSQHEQAHSTGTVGHIPSTYQAOSSQHGIVPESVVAQGSQGG 660
 QY 721 QPSSSILTVSSSQPIQHPPQOQGIQOTAPPOQTVOYSLSTQTSSEATTQAPVOPAP 780
 DB 661 QPSSSILTVSSSQPIQHPPQOQGIQOTAPPOQTVOYSLSTQTSSEATTQAPVOPAP 720
 QY 781 QVLPQVSAKGOSTQGVSAVAPAEPAVAQAPOATQPTTLASSVDSASHSVASGMDGEMV 840
 DB 721 QVLPQVSAKGOSTQGVSAVAPAEPAVAQAPOATQPTTLASSVDSASHSVASGMDGEMV 780
 QY 841 PSSSGRHEGRTTKKHYKSVSRSRHEKTSRPKIRILNVSNKGRVVECOLETHNRKMYT 900
 DB 781 PSSSGRHEGRTTKKHYKSVSRSRHEKTSRPKIRILNVSNKGRVVECOLETHNRKMYT 840
 QY 901 FKPLDGNPEIKITIMWNPFIILAIERESFVDVREIIEKADEMLSDVSVPEBGDGL 960
 DB 841 FKPLDGNPEIKITIMWNPFIILAIERESFVDVREIIEKADEMLSDVSVPEBGDGL 900
 QY 961 ESLOGKDYGFSGSQKLEGEFKQPIPASMPQOIGIPTSSLTQVHVSAGRRFIVSPVES 1020
 DB 901 ESLOGKDYGFSGSQKLEGEFKQPIPASMPQOIGIPTSSLTQVHVSAGRRFIVSPVES 960
 QY 1021 RLRSKVPFSEIITDTVAASTAQSPQMLSHSASSLSLQOAFSELRRAOMTEGNTAPNE 1080
 DB 961 RLRSKVPFSEIITDTVAASTAQSPQMLSHSASSLSLQOAFSELRRAOMTEGNTAPNE 1020
 QY 1081 SHGTGTFPVVPPPLSSISAGVPTTAATAATPVPATSPENDISTSVYQSEVTPTEGLAGV 1140
 DB 1021 SHGTGTFPVVPPPLSSISAGVPTTAATAATPVPATSPENDISTSVYQSEVTPTEGLAGV 1080
 QY 1141 ATSTGVTVSGGLPIPVSESPVLSVSSITIPAVVSISSPQLQVPTSEIIVSSSTA 1200
 DB 1081 ATSTGVTVSGGLPIPVSESPVLSVSSITIPAVVSISSPQLQVPTSEIIVSSSTA 1140
 QY 1201 LYPSTVVSATISASAGSTATPGPKPAPVAVSQAAAGSTTVGATLTVSVTTTSPFSTASQLS 1260
 DB 1141 LYPSTVVSATISASAGSTATPGPKPAPVAVSQAAAGSTTVGATLTVSVTTTSPFSTASQLS 1200
 QY 1261 IQLSSSTSTPTLAEVTVVVSASLDKTSSTTGLAFSLSAPSSSSSSPAGVSSYSYQPG 1320
 DB 1201 IQLSSSTSTPTLAEVTVVVSASLDKTSSTTGLAFSLSAPSSSSSSPAGVSSYSYQPG 1260
 QY 1321 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1380
 DB 1261 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1320

DB 1233 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1292
 QY 1381 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1340
 DB 1293 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1250
 QY 1441 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1500
 DB 1353 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1310
 QY 1501 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1560
 DB 1413 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1370
 QY 1561 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1620
 DB 1473 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1430
 QY 1621 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1680
 DB 1533 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1490
 QY 1681 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1740
 DB 1593 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1550
 QY 1741 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1800
 DB 1653 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1610
 QY 1801 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1860
 DB 1713 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1670
 QY 1861 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1920
 DB 1773 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1730
 QY 1921 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1980
 DB 1833 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1790
 QY 1981 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 2040
 DB 1893 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1850
 QY 2041 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 2100
 DB 1953 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 1910
 QY 2101 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 2160
 DB 2013 LHPVIPSIVIASTPILPOAAGFTSTPILPOVPSIPPIVQPVANVPAVQOTLIHQOPAL 2070

RESULT 14
 ABO44408
 ID ABO44408 standard; protein: 2141 AA.

ABO44408;
 26-SEP-2003 (first entry)

Novel human protein kinase #28.
 Human; kinase; enzyme; cosmetic application; nutraceutical application.

OS Homo sapiens.
 XX US6541252-B1.
 XX 01-APR-2003.

XX 14-MAY-2001; 2001US-00854856.
 XX 19-MAY-2000; 2000US-0206015P.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Hilbun E, Donoho G, Turner CA;
 XX MPI; 2003-575927/54.
 XX N-PSDB; ACH03791.
 XX
 XX New nucleic acid encoding novel human proteins, useful in cosmetic and
 XX nutritional applications.
 XX
 XX Disclosure; Page; 11pp; English.
 XX
 XX The invention relates to a new isolated nucleic acid encoding a novel
 XX human protein kinase. The nucleic acid is useful in cosmetic and
 XX nutritional applications. The present sequence represents the amino
 XX acid sequence of a novel human protein kinase. Note: The sequence data
 XX for this patent did not form part of the printed specification but was
 XX obtained in electronic format directly from USPTO at
 XX seqdata.uspto.gov/sequence.html?docid=654125281
 XX
 XX Sequence 2141 AA;
 XX
 XX Query Match 95.1%; Score 10287.5; DB 7; Length 2141;
 XX Best Local Similarity 94.4%; Pred. No. 0;
 XX Matches 2048; Conservative 0; Mismatches 0; Indels 121; Gaps 2;
 XX
 QY 61 MDKSRGAAATTTTTHRRFRFRSVICDSNATALELPLGLPLSLPOPSPIPAAPQAPPEPH 120
 Db 1 MDKSRGAAATTTTTHRRFRFRSVICDSNATALELPLGLPLSLPOPSPIPAAPQAPPEPH 60
 QY 121 REEVTATATISQVAAQPPAAAPGBOAVAGPAPSTVPSSTSKDNPVSQPSIVSGKEEPP 180
 Db 61 REEVTATATISQVAAQPPAAAPGBOAVAGPAPSTVPSSTSKDNPVSQPSIVSGKEEPP 120
 QY 181 ARSSGGGSAKEPBEBSQDDIELEETKAVGMSNDRLFKPFIETGRSGFKTVYGLD 240
 Db 121 ARSSGGGSAKEPBEBSQDDIELEETKAVGMSNDRLFKPFIETGRSGFKTVYGLD 180
 QY 241 TETTVAVAMCELQDRKLTSEKORFKEBAEMLKGLQHPNIVRFYDSMESTYKGGKCIYV 300
 Db 181 TETTVAVAMCELQDRKLTSEKORFKEBAEMLKGLQHPNIVRFYDSMESTYKGGKCIYV 240
 QY 301 TELMTSGTLTKYLRFPYMKIKVLRSWCROILKGLQFLHTTPTPIIHRDLKCDNIFITGP 360
 Db 241 TELMTSGTLTKYLRFPYMKIKVLRSWCROILKGLQFLHTTPTPIIHRDLKCDNIFITGP 300
 QY 361 TGSVKIGDLGLATLKRPASPAKSVIGTPEFMAPEMYEKEKYDSVVVAFGCMLEMATSEY 420
 Db 301 TGSVKIGDLGLATLKRPASPAKSVIGTPEFMAPEMYEKEKYDSVVVAFGCMLEMATSEY 360
 QY 421 PYSECQNAAOIYRRVTSGVKRPASFDKVAIPEVKEIIEGCIRONKDERYSIKDLINHAFFQ 480
 Db 361 PYSECQNAAOIYRRVTSGVKRPASFDKVAIPEVKEIIEGCIRONKDERYSIKDLINHAFFQ 420
 QY 481 EETGVARELAEDDEGEKIAIKLWRIEDIKLKKGKYKONNEAIEGSPDLERVPEDVAQEM 540
 Db 421 EETGVARELAEDDEGEKIAIKLWRIEDIKLKKGKYKONNEAIEGSPDLERVPEDVAQEM 480
 QY 541 VESGVYEGDHKTAKAIKIDVSLIKRRREOROLVREEOEKKKEESSIKQOVQSSASQ 600
 Db 481 VESGVYEGDHKTAKAIKIDVSLIKRRREOROLVREEOEKKKEESSIKQOVQSSASQ 540
 QY 601 TGIKOLPSASTGIPJASTTSASVSTOVEPREPADQHOLOQYQOPSTISVLDGTVDSDGQ 660
 Db 541 TGIKOLPSASTGIPJASTTSASVSTOVEPREPADQHOLOQYQOPSTISVLDGTVDSDGQ 600
 QY 661 SSVFTESRVSSQOTVTSVGSQHEQAHSTGTVPGHIPSTVQAOSQPHGVYPPSSVAQGSQ 720
 Db 661 SSVFTESRVSSQOTVTSVGSQHEQAHSTGTVPGHIPSTVQAOSQPHGVYPPSSVAQGSQ 720

Db 601 SSVFTESRVSSQOTVTSVGSQHEQAHSTGTVPGHIPSTVQAOSQPHGVYPPSSV----- 653
 QY 721 QPSSSSLTGVSSSQPIQHPQQQGGIQQTAPPOQTVOYISLSQTSSTSSSEATTAQVSPQAP 780
 Db 654 -----QQGICQTAPPOQTVOYISLSQTSSTSSSEATTAQVSPQAP 692
 QY 781 QVLPQVSAGKQ----- 791
 Db 693 QVLPQVSAGKQFPPEPLPPQYPGDSNTAPSSSNVASCIIHSTVYLKPEMTEVLATPGYEPT 752
 QY 792 -----STQGVQVAPAPAV 807
 Db 753 VVQPVESNLLVPMGVGVGVQVQPGGSLAQAPTSTSSQAVLESTQGVQVAPAPAV 812
 QY 808 AQPOATQPTTLASSVDSAHSDVAGMSDGNENVPSSGRHGGRTTKRHYKSVRSRSH 867
 Db 813 AQPOATQPTTLASSVDSAHSDVAGMSDGNENVPSSGRHGGRTTKRHYKSVRSRSH 872
 QY 868 KTSRPKRLINVSNKGDVAVBCOLETHNRKMWTFKPDLDGDNPEIATIMVNNPILAE 927
 Db 873 KTSRPKRLINVSNKGDVAVBCOLETHNRKMWTFKPDLDGDNPEIATIMVNNPILAE 932
 QY 928 RESFVDQVRELIIEKADDEMLSEDVSVBEPEGDLGLSLQKDDYGRSGQKLEGEFKOPIPA 987
 Db 933 RESFVDQVRELIIEKADDEMLSEDVSVBEPEGDLGLSLQKDDYGRSGQKLEGEFKOPIPA 992
 QY 988 SSMPOQIGIPITSSLTQVHSGRRFIVSPVESLRSKVPSEITTDVAASTQSPGM 1047
 Db 993 SSMPOQIGIPITSSLTQVHSGRRFIVSPVESLRSKVPSEITTDVAASTQSPGM 1052
 QY 1048 LSHSASSLSLQQAASELRRAQMTGEPNTAPNFBHTGTFPVPVPELSSIAIGVPTTAAT 1107
 Db 1053 LSHSASSLSLQQAASELRRAQMTGEPNTAPNFBHTGTFPVPVPELSSIAIGVPTTAAT 1112
 QY 1108 APVATSSPNDISTSVIYSEVTVPTTEGIAAGVATSTGVTSQGLPIPVSESPVLSVV 1167
 Db 1113 APVATSSPNDISTSVIYSEVTVPTTEGIAAGVATSTGVTSQGLPIPVSESPVLSVV 1172
 QY 1168 SSTIIPAVSISTTSPBLQVPTSTSEIVSTALYPSVTYATASAGSITATGPRPA 1227
 Db 1173 SSTIIPAVSISTTSPBLQVPTSTSEIVSTALYPSVTYATASAGSITATGPRPA 1232
 QY 1228 VVSQQAAGSTTVAGTLTSTVSTTSPSTASQSLSTLSSSTPTLAEVTVVVSASLDTTS 1287
 Db 1233 VVSQQAAGSTTVAGTLTSTVSTTSPSTASQSLSTLSSSTPTLAEVTVVVSASLDTTS 1292
 QY 1288 HSSITGLAFSLASBSSSSPGAGVSYISQPGIHLPIVBSVIASTPILPQAAGPTSTPL 1347
 Db 1293 HSSITGLAFSLASBSSSSPGAGVSYISQPGIHLPIVBSVIASTPILPQAAGPTSTPL 1352
 QY 1348 LPQVPSIPPLVOPAVNPVAVOQTLHSQOPALLPNPHHCPEVVDSTQKAGIDIK 1407
 Db 1353 LPQVPSIPPLVOPAVNPVAVOQTLHSQOPALLPNPHHCPEVVDSTQKAGIDIK 1412
 QY 1408 TLEBKLSLSEHSSSGAQAASVLSLETSVLESTVGTGIPITTAAPSKLTLSTSTCLP 1467
 Db 1413 TLEBKLSLSEHSSSGAQAASVLSLETSVLESTVGTGIPITTAAPSKLTLSTSTCLP 1472
 QY 1468 TNLPLGTVALPVPVTPGVSTPVSTTSGVKRGTAPSKPELTKAPVLPVTELPAGTL 1527
 Db 1473 TNLPLGTVALPVPVTPGVSTPVSTTSGVKRGTAPSKPELTKAPVLPVTELPAGTL 1532
 QY 1528 PSEQLPPEPSPSLTQSQOPELDDAQLRKLSPKXITVTSVAVGVSMALPALTAEATOP 1587
 Db 1533 PSEQLPPEPSPSLTQSQOPELDDAQLRKLSPKXITVTSVAVGVSMALPALTAEATOP 1592
 QY 1588 QKGVSYKESPVLATSSGAVFMKRGFQVVAADGAKESKNSSEDAKSHFESSSTESS 1647
 Db 1593 QKGVSYKESPVLATSSGAVFMKRGFQVVAADGAKESKNSSEDAKSHFESSSTESS 1652
 QY 1648 VLSSSSESTLVKPEPNGITIPGISDVPESAHKTASAKSDPTGQPTKVGRRPVTTAN 1707
 Db 1653 VLSSSSESTLVKPEPNGITIPGISDVPESAHKTASAKSDPTGQPTKVGRRPVTTAN 1712

QY 1708 KVGKFSVSKTEDEKTTDKKEGVPVAPPEMDLEQAVLPAVLPKKEKPELSEPSHINNGSSD 1767
 DB 1713 KVGKFSVSKTEDEKTTDKKEGVPVAPPEMDLEQAVLPAVLPKKEKPELSEPSHINNGSSD 1772
 QY 1768 PEAAPFLSRDVEDDGSPPHSPHQLSSKSLPSQNLSSNSFNSSSYMSDNESEDIEDDLK 1827
 DB 1773 PEAAPFLSRDVEDDGSPPHSPHQLSSKSLPSQNLSSNSFNSSSYMSDNESEDIEDDLK 1832
 QY 1828 LELRLRLDKHLKEIODIQRKHEIESLYTKLGKVPVAVIIPPAAPLSGRRRRRTKSGS 1887
 DB 1833 LELRLRLDKHLKEIODIQRKHEIESLYTKLGKVPVAVIIPPAAPLSGRRRRRTKSGS 1892
 QY 1888 KSSRSSSLGNKSPQLSGNLSSGSAAYLHPQOTLHPGNIPESSQNLLOPLKSPSSDN 1947
 DB 1893 KSSRSSSLGNKSPQLSGNLSSGSAAYLHPQOTLHPGNIPESSQNLLOPLKSPSSDN 1952
 QY 1948 LYSAPFSDGAISSVPSLSAPGCGTSTNTVGTATVNSQAQAOPPAMTSSRKCTFTDDHLK 2007
 DB 1953 LYSAPFSDGAISSVPSLSAPGCGTSTNTVGTATVNSQAQAOPPAMTSSRKCTFTDDHLK 2012
 QY 2008 VDNARADAMNLSGRGSKGHNVGPGMARKEFSAPOULCISMTNLSGSAFISAASATSL 2067
 DB 2013 VDNARADAMNLSGRGSKGHNVGPGMARKEFSAPOULCISMTNLSGSAFISAASATSL 2072
 QY 2068 GHFTRSMCPPOQYGFPAATPFGAOWSGTGAPAPQPLGQFQPVGTASLQNFNISNLQKSISSN 2127
 DB 2073 GHFTRSMCPPOQYGFPAATPFGAOWSGTGAPAPQPLGQFQPVGTASLQNFNISNLQKSISSN 2132
 QY 2128 PPGSNLRTT 2136
 DB 2133 PPGSNLRTT 2141

RESULT 15

ABO4405

ID ABO4405 standard; protein, 2294 AA.

AC ABO4405;

DT 26-SEP-2003 (first entry)

DE Novel human protein kinase #25.

KW Human; kinase; enzyme; cosmetic application; nutraceutical application.

OS Homo sapiens.

PN US6541252-B1.

PD 01-APR-2003.

PF 14-MAY-2001; 2001US-00854856.

PR 19-MAY-2000; 2000US-0206015P.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Hilbun E, Donoho G, Turner CA;

DR WPI; 2003-575927/54.

DR N-PSDB; ACH03788.

PT New nucleic acid encoding novel human proteins, useful in cosmetic and nutraceutical applications.

PS Disclosure; Page; 11pp; English.

CC The invention relates to a new isolated nucleic acid encoding a novel human protein kinase. The nucleic acid is useful in cosmetic and nutraceutical applications. The present sequence represents the amino acid sequence of a novel human protein kinase. Note: The sequence data for this patent did not from part of the printed specification but was

CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=6541252b1
 XX
 SQ Sequence 2294 AA;

Query Match 94.4%; Score 10211; DB 7; Length 2294;
 Best Local Similarity 88.2%; Pred. No. 0;
 Matches 2048; Conservative 0; Mismatches 0; Indels 274; Gaps 2;

QY 61 MDKDSRAAATTTTTHRRFRFRSVICDSNATALEPGLPISLPQPSIPAAVQSAAPPBP 120
 DB 1 MDKDSRAAATTTTTHRRFRFRSVICDSNATALEPGLPISLPQPSIPAAVQSAAPPBP 60
 QY 121 REETVATATSOVAQAPPAAPAPGQAVAPAPSTVSSTKORPVSOPIVSGKEPPP 180
 DB 61 REETVATATSOVAQAPPAAPAPGQAVAPAPSTVSSTKORPVSOPIVSGKEPPP 120
 QY 181 ARSGSGGSAKEPEEESQOQDDIELEETAVAGMSNDRFLKPIEIGRGSFKTVYGLD 240
 DB 121 ARSGSGGSAKEPEEESQOQDDIELEETAVAGMSNDRFLKPIEIGRGSFKTVYGLD 180
 QY 241 TETTVAVMCELODRKLTSPRORFKEAEMLKGLQHPNIVRFYDSMESTYKAKCIYLV 300
 DB 181 TETTVAVMCELODRKLTSPRORFKEAEMLKGLQHPNIVRFYDSMESTYKAKCIYLV 240
 QY 241 TELMTSGTLKTYLKRFRVWKIKVLRWCROILKGLQELHRTTPEIIRHDLKDNIFITGP 360
 DB 241 TELMTSGTLKTYLKRFRVWKIKVLRWCROILKGLQELHRTTPEIIRHDLKDNIFITGP 300
 QY 361 TGSVKIGDLGIATIKRASPAKSVIGTEPFMAPEYKEKYDESUVYAFGMCMLEATSEY 420
 DB 301 TGSVKIGDLGIATIKRASPAKSVIGTEPFMAPEYKEKYDESUVYAFGMCMLEATSEY 360
 QY 421 PYSECQNAAOIYRVTSGVWPASFDKVAIPVKEIIEGCTIRONDERYSIDLNHAFQ 480
 DB 361 PYSECQNAAOIYRVTSGVWPASFDKVAIPVKEIIEGCTIRONDERYSIDLNHAFQ 420
 QY 481 EETGVAVELAEEDDEGEKIAIKMLRIDIKKLGKQYKQNEAIEFSPLERDVEDVAQEM 540
 DB 421 EETGVAVELAEEDDEGEKIAIKMLRIDIKKLGKQYKQNEAIEFSPLERDVEDVAQEM 480
 QY 541 VESGVYCBGDKHTMAKAIKDRVSLIKRREORQVREBQEKKEBESLKQCVQSSASQ 600
 DB 481 VESGVYCBGDKHTMAKAIKDRVSLIKRREORQVREBQEKKEBESLKQCVQSSASQ 540
 QY 601 TGIKOLPSASNGIPTASTASVSTOVEPEEPADQHQLOLYOQPSISVLSGTVDSGQ 660
 DB 541 TGIKOLPSASNGIPTASTASVSTOVEPEEPADQHQLOLYOQPSISVLSGTVDSGQ 600
 QY 661 SSVFTESRVSSQQTVSYSQHEQASHGTGVBGHIPTVQAQSGPHGVYPPSSVAQGSQ 720
 DB 601 SSVFTESRVSSQQTVSYSQHEQASHGTGVBGHIPTVQAQSGPHGVYPPSSVAQGSQ 660
 QY 721 QPSSSSLTGVSSSQPIQHPQOQOGIQTAPRQOQVYVSLSTSSSEKTTAPVSOQAP 780
 DB 654 QPSSSSLTGVSSSQPIQHPQOQOGIQTAPRQOQVYVSLSTSSSEKTTAPVSOQAP 720
 QY 781 QVLPQVSAKQ-----QQGIQQTAPQOQVYVSLSTSSSEKTTAPVSOQAP 791
 DB 693 QVLPQVSAKQLPVSAQVPVPTIGBPQLPVATQPSVVPVHSAHFLPVGQPLPLPQXP 752
 QY 792 ----- 791
 DB 753 VSGIPISRPVHSTAGTGFSSLPITMAAGITQPLTLASSATTAIPGVSTVPSQLPTLL 812
 QY 792 ----- 791
 DB 813 QPVTQLPEQVHPOLLQPAVQSMGIPANLGOAAEVPLSSGDLVLYGFPPLPPOYGDGSI 872
 QY 792 ----- 791
 DB 873 APSSVAVSVCHSTVLAHPMPTEVLATGTFPTVQPVVESNLVPMQGVGVGVQVQSG 932

QY 792 -----STQGVQVAPAPVPAVPAOQATQPTTLASVDYSAHSDVAGS 834
 Db 993 GSIAQAPITSSQOAVLESTQGVQVAPAPVPAVPAOQATQPTTLASVDYSAHSDVAGS 992
 QY 895 DGENAVPSSSGHBERITTKRHVRKSVRSRREKTSRPRKLILVSNKGRVVECOLETH 894
 Db 993 DGENAVPSSSGHBERITTKRHVRKSVRSRREKTSRPRKLILVSNKGRVVECOLETH 1052
 QY 895 NRKMTFFEDLDGNDPEEIIATIMVANDFIATERESFVQVVEIIEKADENISEVSVSE 954
 Db 1053 NRKMTFFEDLDGNDPEEIIATIMVANDFIATERESFVQVVEIIEKADENISEVSVSE 1112
 QY 955 EGDQGLJSLQKDDYGFSGSQKLEGEFKQPIPASSMPOQIGIPISLTQVHSAGRFIV 1014
 Db 1113 EGDQGLJSLQKDDYGFSGSQKLEGEFKQPIPASSMPOQIGIPISLTQVHSAGRFIV 1172
 QY 1015 SPVPSRRLRESKVFSEIITDYVAASTAOSPQNLSSHSSLSLQOAFSELRAQMTSEPN 1074
 Db 1173 SPVPSRRLRESKVFSEIITDYVAASTAOSPQNLSSHSSLSLQOAFSELRAQMTSEPN 1232
 QY 1075 TAPNFSHTGTFPVVPPFLSSIAGVPTTAATAPVPATSSPNDISTSVIOSEVTVTE 1134
 Db 1233 TAPNFSHTGTFPVVPPFLSSIAGVPTTAATAPVPATSSPNDISTSVIOSEVTVTE 1292
 QY 1135 EGIAGVATSTGVVTSGLPIPEVSEPVLSVSSITIPAVVSIITTSPLSQVPTSTSEI 1194
 Db 1293 EGIAGVATSTGVVTSGLPIPEVSEPVLSVSSITIPAVVSIITTSPLSQVPTSTSEI 1352
 QY 1195 VVSSTALYPSVTGATSSAGSSTRATBPKPPAVVSOQAAGSTTVGATLTSTVSTTSFPS 1254
 Db 1353 VVSSTALYPSVTGATSSAGSSTRATBPKPPAVVSOQAAGSTTVGATLTSTVSTTSFPS 1412
 QY 1255 TASQLSIOLSSSTSTPTLAETVWVSAHSLDKTSHSTGLAFSLJAPSSSSSPGAGVSSY 1314
 Db 1413 TASQLSIOLSSSTSTPTLAETVWVSAHSLDKTSHSTGLAFSLJAPSSSSSPGAGVSSY 1472
 QY 1315 ISQPGHLPVTPSVIASTPILPQAAGTSTPLLPQVPSIPPLVQPVANVPVAVOQTLIHS 1374
 Db 1473 ISQPGHLPVTPSVIASTPILPQAAGTSTPLLPQVPSIPPLVQPVANVPVAVOQTLIHS 1532
 QY 1375 QOPQALLPNQPHTCPEVSDTQPKAPGIDIKITLEKRLSEHSSSGAQAHSVSIET 1434
 Db 1533 QOPQALLPNQPHTCPEVSDTQPKAPGIDIKITLEKRLSEHSSSGAQAHSVSIET 1592
 QY 1435 SLVSESTVTPGIPPTAIVAPSKLTSTSTCLPPTNLPGTVALPVTVPYTPGOVSTPVST 1494
 Db 1593 SLVSESTVTPGIPPTAIVAPSKLTSTSTCLPPTNLPGTVALPVTVPYTPGOVSTPVST 1652
 QY 1495 TTSGVKFGTAPSKPPLTKAPVLPVGTLPAGTLPSEQLPFPGPSLTQSQOPLLEDLDAQL 1554
 Db 1653 TTSGVKFGTAPSKPPLTKAPVLPVGTLPAGTLPSEQLPFPGPSLTQSQOPLLEDLDAQL 1712
 QY 1555 RRTLSPEXITVTSVAVGVMAAPTAITBAGTQPOKGVQVKEGVLATSSGAGVFKMGGRF 1614
 Db 1713 RRTLSPEXITVTSVAVGVMAAPTAITBAGTQPOKGVQVKEGVLATSSGAGVFKMGGRF 1772
 QY 1615 QVSVAADGAQKEGKNKSEDAKSVFESSTSESSVLSSSSPSTLVKPEPNGITIPGISD 1674
 Db 1773 QVSVAADGAQKEGKNKSEDAKSVFESSTSESSVLSSSSPSTLVKPEPNGITIPGISD 1832
 QY 1675 VPESAHKTTASAEKSDTGQPTKVGRFOVTTTANKVGRFVSXKTEBKITDTKKEGPVAPSP 1734
 Db 1833 VPESAHKTTASAEKSDTGQPTKVGRFOVTTTANKVGRFVSXKTEBKITDTKKEGPVAPSP 1892
 QY 1735 FMDLEQAVLPAVLPKKEKPELSESHLNGPSDPEAFLSRDVUDGSGSPHSPHQLSSKS 1794
 Db 1893 FMDLEQAVLPAVLPKKEKPELSESHLNGPSDPEAFLSRDVUDGSGSPHSPHQLSSKS 1952
 QY 1795 LPSQNTLSQSLNSFNSSYMSDNEJIEDEDLKLELRDLKHLKEIODLOSROKHEIES 1854
 Db 1953 LPSQNTLSQSLNSFNSSYMSDNEJIEDEDLKLELRDLKHLKEIODLOSROKHEIES 2012
 QY 1855 LYTGLGVPPAVIIPPAAPLGGRRRPTKSGSKSRSSSLGNKSPQLSGNTLSGQSAASY 1914

Db 2013 LYTGLGVPPAVIIPPAAPLGGRRRPTKSGSKSRSSSLGNKSPQLSGNTLSGQSAASY 2072
 QY 1915 LHPQOTLHPGNIPESGQNOLOPLKPSSSDNLYSAFTSDGALSVPLSAPGQGTSSSTN 1974
 Db 2073 LHPQOTLHPGNIPESGQNOLOPLKPSSSDNLYSAFTSDGALSVPLSAPGQGTSSSTN 2132
 QY 1975 TVGATVNSQAAQAPPAATSSRKGTFTDHLKLVNWARDAMNLSGRGSGKHNNYEGPG 2034
 Db 2133 TVGATVNSQAAQAPPAATSSRKGTFTDHLKLVNWARDAMNLSGRGSGKHNNYEGPG 2192
 QY 2035 MARKFSAPGOLCISNTSNLGGSAPIASAATSLGHFTSMCPPOQYGFPAATPFGAOWSGT 2094
 Db 2193 MARKFSAPGOLCISNTSNLGGSAPIASAATSLGHFTSMCPPOQYGFPAATPFGAOWSGT 2252
 QY 2095 GGPAPQPLGQFQPVGTASLQNFNINSLQKSIISNPPGSLRFTT 2136
 Db 2253 GGPAPQPLGQFQPVGTASLQNFNINSLQKSIISNPPGSLRFTT 2294

Search completed: September 24, 2004, 01:06:44
 Job time : 213 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 25, 2004, 05:24:04 ; Search time 1217 Seconds
(without alignments)
7380.239 Million cell updates/sec

Title: US-10-010-720-14
Perfect score: 10812
Sequence: 1 MSGAAMEKOSTPGSLFLSP.....NISNLQKSIISPPGSNLRTT 2136

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6597448 seqs, 2102469792 residues
Total number of hits satisfying chosen parameters: 13194896

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPRO_spool/US10010720/runat_23092004_154321_25331/app_query.fasta.1.2311
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rpnp -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt
-LIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MOB=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10010720 @CGN 1.1 249 @runat 23092004 154321 25331 -NCPU=6 -ICPU=3
-NO MMAP -LARGEIOBRY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA New :
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq :
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq :
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq :
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq :
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq :
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq :
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq :
8: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq :
9: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq :

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	10677	98.8	7149 1	PCT-US04-24424-2014
2	10677	98.8	7149 1	PCT-US04-27459-1
3	10646	98.5	7152 7	US-10-491-467-54
4	10604	98.1	8678 7	PCT-US04-02652-450
5	10604	98.1	8678 7	US-10-767-471-450
6	10604	98.1	10464 1	PCT-US04-02652-455
7	10604	98.1	10464 7	US-10-767-471-455
8	10567.5	97.7	7853 1	PCT-US04-02652-452
9	10567.5	97.7	7853 7	US-10-767-471-452

10	10567.5	97.7	9639 1	PCT-US04-02652-454	Sequence 454, App
11	10567.5	97.7	9639 7	US-10-767-471-454	Sequence 454, App
12	9737	90.1	9417 7	PCT-US04-02652-449	Sequence 449, App
13	9737	90.1	9417 7	US-10-767-471-449	Sequence 449, App
14	9729.5	90.0	7328 5	US-09-979-167-63	Sequence 63, App
15	7154	66.2	4872 1	PCT-US04-02652-453	Sequence 453, App
16	7154	66.2	4872 7	US-10-767-471-453	Sequence 453, App
17	5931	54.9	5765 7	US-10-021-6984-3606	Sequence 3606, App
18	5902	54.6	6294 1	PCT-US04-02652-451	Sequence 451, App
19	5902	54.6	6294 7	US-10-767-471-451	Sequence 451, App
20	3132.5	29.0	170546 1	PCT-US04-02652-10703	Sequence 10703, A
21	3132.5	29.0	170546 7	US-10-767-471-10703	Sequence 10703, A
22	3026	28.0	3835 6	US-10-765-700-1120	Sequence 120, App
23	2719.5	25.2	5637 6	US-10-932-349-491	Sequence 491, App
24	2642	24.4	7825 7	US-10-490-592-6	Sequence 4, App
25	2639.5	24.4	6582 7	US-10-490-592-6	Sequence 6, App
26	2638.5	24.4	6891 1	PCT-US04-27459-5	Sequence 5, App
27	2627	24.3	1536 5	US-09-552-317-7020	Sequence 7020, App
28	2623	24.3	6834 1	PCT-US04-27459-3	Sequence 3, App
29	2526.5	23.4	7569 6	US-10-932-349-494	Sequence 494, App
30	2466	22.8	5403 1	PCT-US04-27459-6	Sequence 6, App
31	2441.5	21.6	7278 1	PCT-US04-27459-8	Sequence 8, App
32	2355.5	21.6	6807 1	PCT-US04-27459-7	Sequence 7, App
33	2130	19.7	3861 1	PCT-US04-27459-9	Sequence 9, App
34	2096.5	19.4	4064 6	US-10-932-349-356	Sequence 356, App
35	2096.5	19.4	4341 6	US-10-932-349-357	Sequence 357, App
36	1974	18.3	1928 1	PCT-US04-27459-2	Sequence 2, App
37	1922	17.8	1345 7	US-10-021-6984-3609	Sequence 3609, App
38	1807	16.7	2404 1	PCT-US04-27459-4	Sequence 4, App
39	1802	16.7	2328 6	US-10-932-349-493	Sequence 493, App
40	1718	15.9	1667 7	US-10-100-683-2951	Sequence 2951, App
41	1508	13.9	1179 7	US-10-021-6984-3607	Sequence 3607, App
42	1444	13.4	893 6	US-10-472-665-141	Sequence 141, App
43	1444	13.4	893 6	US-10-896-164-1520	Sequence 1520, App
44	1444	13.4	893 6	US-10-868-194B-1520	Sequence 1520, App
45	1444	13.4	893 6	US-10-868-184B-1520	Sequence 1520, App

ALIGNMENTS

RESULT 1
PCT-US04-24424-2014
; Sequence 2014, Application PC/TUS0424424
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rudnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PCT
; CURRENT APPLICATION NUMBER: PCT/US04/24424
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US 60/490, 890
; PRIOR FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2786
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2014
; LENGTH: 7149
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-24424-2014

Alignment Scores:

Pred. No.: 0
Score: 10677.00
Percent Similarity: 89.63%
Best Local Similarity: 89.63%
Query Match: 98.75%
DB: 1
Gaps: 1

US-10-010-720-14 (1-2136) x PCT-US04-24424-2014 (1-7149)
Length: 7149
Matches: 2135
Conservative: 0
Mismatch: 1
Indels: 246

```
QY 1 MetSerGIYGIYAlaIaIaGIuIuYSerSerThrProGIYSerLeuPheLeuSerPro 20
Db 1 ATGCTGGCGCGCGCGAGAGAGACAGACACTCCCGGTTCCCTGCTTCCGCG 60
QY 21 ProAlaProAlaProLYaAsnGIYSerSerSerSerSerValGIYGIuLYeLeuGIY 40
Db 61 CCGGCTCCTGGCCCCCAAGAAATGGCTCCAGCTCCGATCTCCGTGGGGGAGAAACTGGGA 120
QY 41 AlaAlaAlaAlaAspAlaValThrGIYArgThrGIuGIuTYrArgArgArgArgHisThr 60
Db 121 GCCCGCGCGCGCGAGCTGTGTACCGGACGAGACGAGAGTACAGCGCGCGCGCACACT 180
QY 61 MetAspLYaAspSerArgLYAlaAlaAlaThrThrThrThrThrGluHisArgPhePhe 80
Db 181 ATGACACAAGACACACCGTGGGGCGCGCGCACACTACACCACTGACACCGCTTCTTC 240
QY 81 ArgArgSerValIleCYaAspSerAsnAlaThrAlaLeuGIuLeuProGIYLeuProLeu 100
Db 241 CGCGGAGCGCTCATCTGGACTCCCAATGCCACTGGCTGGAGCTTCCGCGGCTTCTCTCT 300
QY 101 SerLeuProGIuProSerIleProAlaAlaValProGIuSerAlaProProGIuProHis 120
Db 301 TCCCTGCCCGCGCGACAGATCCCGCGGTGTCCGACAGTGTCCACCGAGCGCCAC 360
QY 121 ArgGIuGIuThrValThrAlaThrAlaThrSerGIuValAlaGIuGIuProProAlaAla 140
Db 361 CGGAGAGAGACCGTGAACCGCCACCTTCCAGGTAGCCGACCGACCTCCAGCGCT 420
QY 141 AlaAlaProGIYGIuGIuAlaValAlaGIYProAlaProSerThrValProSerSerThr 160
Db 421 GCCCGCCCCGGGAGAACAGCGCTGGCGGCGCTCCCTCCGACTGTGCCAGAGTACC 480
QY 161 SerLYaAspArgProValSerGIuProSerLeuValGIYSerLYeGIuGIuProProPro 180
Db 481 AGCAAAAGCGCGCGAGTGTCCAGCTTGTGTGGAGCAAAAGAGAGAGCGCGCGCG 540
QY 181 AlaArgSerGIYSerGIYGIYSerAlaLYeGIuGIuProGIuGIuGIuArgSerGIuGIu 200
Db 541 GCGAGAGTGGCAACGGCGCGCGCGCGCGCGCACAGACACAGAGAGAGAGAGAGAG 600
QY 201 GIuAspAspIleGIuGIuLeuGIuThrLYeAlaValGIYMetSerAsnAspGIYArgPhe 220
Db 601 CAGATATATATCGAAGAGCTGAGACCAAGCGCTGGGAGATGTAAAGATGGCGCTT 660
QY 221 LeuLYePheAspIleGIuIleGIYArgGIYSerPheLYeThrValTYrLYeGIYLeuAsp 240
Db 661 CTCAAGTTTGACATCGAATCGGAGGCTCTTTAAGACGTCTCAAAAGGTCTGAC 720
QY 241 ThrGIuThrThrValGIuValAlaTYrCYeGIuLeuGIuAspArgLYeThrLYeSer 260
Db 721 ACTGAAACCAACCGTGAAGTGGCTGGTGTGAATGCGAGGATCGAAAATTAAACAAAGCT 780
QY 261 GIuArgGIuArgPheLYeGIuGIuAlaGIuMetLeuLYeGIYLeuGIuHisProAsnIle 280
Db 781 GAGAGGAGAGATTTAAAGAAAGCTGAATGTAAAGAGTCTTCAGCATCCCAATATT 840
QY 281 ValArgPheTYrAspSerTYrPGLuSerThrValLYeGIYLYeLYeCysIleValLeuVal 300
Db 841 GTTGATTTTATGATCTCTGGGAATCCACGATTAAGAAAGAAAGATGTGTTTGTG 900
QY 301 ThrGIuLeuMetThrSerGIYThrLeuLYeThrTYrLeuLYeArgPheLYeValMetLYe 320
Db 901 ACTCAACTTATGACGTCTGGAACACTTAACGATCTGAAAAGGTTTAAAGATGATGAG 960
QY 321 IleLYeValLeuArgSerTYrCYeArgGIuIleLeuLYeGIYLeuGIuPheLeuHisThr 340
Db 961 ATCAAACTTCAAAAGCTGTGTCCGTCAAGATCTTAAAGGTCTTCGATTTCTTCATACT 1020
QY 341 ArgThrProProIleIleHisArgAspLeuLYeCysAspAsnIlePheIleThrGIYPro 360
Db 1021 CGAAGCTCACCTTATTCACCGGATCTTAAAGTGAACAACATCTTATTCACCGGCGCT 1080
QY 361 ThrGIYSerValLYeIleGIYAspLeuGIYLeuAlaThrLeuLYeArgAlaSerPheAla 380
Db 1081 ACTGCTCAGTCAGATGAGAGACTCGCTGGCAACCTTGAAAGCGGCGCTTCTTTGCC 1140
QY 381 LYeSerValIleGIYThrProGIuPheMetAlaProGIuMetTYrGIuGIuLYeTYrAsp 400
Db 1141 AAGAGTGTATAGTACCCCGAGAGTTGATGAGCGCCCTGAGATGTATGAGAGAAATATGAT 1200
QY 401 GluSerValAspValTYrAlaPheGIYMetCysMetLeuGIuMetAlaThrSerGIYTYr 420
Db 1201 GAATCCGTGACGTTTATGCTTTGGATGTGATGCTTGAAGAGGCTACATCTGAATAT 1260
QY 421 ProTYrSerGIuCYeGIuAsnAlaAlaGIuIleTYrArgArgValAlaThrSerGIYValLYe 440
Db 1261 CTTACTCGAGAGTCCAAATATGCTGCGCAGATCTTACCGTCCGCGACAGTGGGTGAG 1320
QY 441 ProAlaSerPheAspLYeValAlaIleProGIuValLYeGIuIleGIuGIYCysIle 460
Db 1321 CCAGCCAGTTTTCGAAAGTGAATTCCTGAAGTGAAGAAATTAATGAAGATGACATA 1380
QY 461 ArgGIuAsnLYeAspGIuArgTYrSerIleLYeAspLeuLeuAsnHisAlaPhePheGIu 480
Db 1381 CGACAAACAAAGATGAAGATATTCATCAACCACTTTGAAACCATGCTTCTTCCAA 1440
QY 481 GIuGIuThrGIYValArgValGIuLeuAlaGIuGIuAspAspGIYGIuLYeIleAlaIle 500
Db 1441 GAGGAAACAGAGAGTACCGGTATGAAATTAACAGAGAAAGTATGAGAGAAATAATGACATA 1500
QY 501 LYeLeuTYrLeuArgIleGIuAspIleLYeLYeLYeLYeLYeLYeLYeLYeLYeLYe 520
Db 1501 AAATTTATGGCTACGATTTGAAGATATTAAGAAATTAAGGAGAAATPACAAAGATATGAA 1560
QY 521 AlaIleGIuPheSerPheAspLeuGIuArgAspValProGIuAspValAlaGIuGIuMet 540
Db 1561 GCTATTAGATTTCTTTGATTATTAAGAGAGATGTCCAGAGAGTGTGCAAGAAATG 1620
QY 541 ValGIuSerGIYTYrValCYeGIuGIYAspHisLYeThrMetAlaLYeAlaIleLYeAsp 560
Db 1621 GTTAGAGTGGTATGTCTGTGAAGTGTATCAGACAGACATGCTTAAGCTATTAAGAC 1680
QY 561 ArgValSerLeuIleLYeArgLYeArgGIuGIuArgGIuLeuValArgGIuGIuGIu 580
Db 1681 AGAGTATCATTAATTAAGAGAAACGAGACAGCGGAGTGTGAACGGAGAGCAAGAA 1740
QY 581 LYeLYeLYeGIuGIuGIuSerSerLeuLYeGIuGIuAlaGIuGIuSerSerAlaSerGIu 600
Db 1741 AAAAAGAGAGAGAGAGAGAGCTCAAAAGAGAGTGAACATCAAGTCTCCAG 1800
QY 601 ThrGIYIleLYeGIuLeuProSerAlaSerThrGIYIleProThrAlaSerThrThrSer 620
Db 1801 ACAGGAATCAAGACGCTCCCTTGCTAGACACGGGATACCTACTGCTTACACACTTCA 1860
QY 621 AlaSerValSerThrGIuValGIuProGIuGIuArgGIuAlaAspGIuHisGIuGIuLeu 640
Db 1861 GCTTCACTTTCTACACAGATGAACCTGAAGAACTTAGGACAGATCAACATCAACACTA 1920
QY 641 GIuTYrGIuGIuProSerIleSerValLeuSerAspGIYThrValAspSerGIYGIuGIY 660
Db 1921 CAGTACCAAGAACCCCAATATATCTGTATCTGATGGACGGTTGACAGTGGTCAAGGA 1980
QY 661 SerSerValPheThrGIuSerArgValSerSerGIuGIuThrValSerTYrGIYSerGIu 680
Db 1981 TCCCTGTCTTTCACAGATCTCGAGTGACGACGCAACAGACAGTTTATATGTTGCCAA 2040
QY 681 HisGIuGIuAlaHisSerThrGIYThrValProGIYHisIleProSerThrValAlaAla 700
Db 2041 CATGAACAGCAATCTTACAGGACAGTCCAGGGCATATACCTTCTACCTGTCACAGCA 2100
QY 701 GIuSerGIuProHisGIYValTYrProProSerSerValAlaGIuGIuYGIuSerGIuGIY 720
Db 2101 CAGCTTACGCCCAATGGGCTATATTCACCTCAAGTGTGGACAGGGGCAAGGCCAGGCT 2160
QY 721 GIuProSerSerSerLeuThrGIYValSerSerSerGIuProIleGIuHisProGIu 740
```


QY	1215	GLYISERTTHRALATHProGlyProlysProProlaValaValaSerGlnAla	1234
Db	4381	GGGGGAGTACTGCTACCCAGAGCTTACGCTTCAGGTGAGTATCTCACAGAGCGCA	4440
QY	1235	GlySerThrThrValaGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer	1254
Db	4441	GGCAGCACTACTGTTGGAGCCACATTTACATCAGTTCTTACACCACTTCATCCCAAGC	4500
QY	1255	ThrAlaSerGlnLeuSerTlleGlnLeuSerSerSerThrSerThrProthLeuAlaGlu	1274
Db	4501	ACAGCTTACAGCTGTCCTTCATTCAGCTTACAGCAGACTTCTTACTCTTACTTAACTGAA	4560
QY	1275	ThrValValaValaSerAlaHisSerLeuaspPyThrSerHisSerSerThrThrGlyLeu	1294
Db	4561	ACCGGGAGTATTAGGACACACTCACTAGATTAAGATCTCATAGAGTACAACTGGATTG	4620
QY	1285	AlaPheSerLeuSerAlaProSerSerSerSerProGlyAlaGlyValSerSerTyr	1314
Db	4621	GCTTCTCCCTCTCGACCACTTCCTTCCTTCCTTCCTCGAGGAGGAGTGTCAATTAT	4680
QY	1315	IlleSerGlnProGlyGlyLeuHisProLeuValIlleProSerValIlleAlaSerThrPro	1334
Db	4681	ATTTCACGCGCGTGGGGCTGCATCCTTTGGTCATTCATAGTGAATAGCTTACTACTCT	4740
QY	1335	IlleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerTlle	1354
Db	4741	ATTCTTCCCAAGCAGCAGGAGCACTTCTACCTTTATTAACCCCAAGTACCTAAGTATC	4800
QY	1355	ProProLeuValaGlnProValAlaAsnValProAlaValaGlnGlnThrLeuIlleHisSer	1374
Db	4801	CCACCTTGTGATACAGCTCTGTGCCATATGCTGCTGTACAGCAGACACTAATTCATAGT	4860
QY	1375	GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer	1394
Db	4861	CAGCTCAACCAAGCTTGTCTCTCCCAACCAAGCCCATACTCATTTGCTCGAAGTATGATCT	4920
QY	1395	AspThrGlnProLysAlaProGlyTlleAspAspIleLysThrLeuGlnGluLysLeuArg	1414
Db	4921	GATACACAAACCAAGCTCTCGAATATATACATAAAGACTCTAAGAAAGAAACCTGGG	4980
QY	1415	SerLeuPheSerGluHisSerSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThr	1434
Db	4981	TCTCGTTCACTGAACACAGCTCATCTGAGCTCAGCATGCTCTGTCTACTGGAGACC	5040
QY	1435	SerLeuValIlleGluSerThrValThrProGlyTlleProThrAlaValaAlaProSer	1454
Db	5041	TCACCTAGCATAGAGAGACGTGTACACACAGCACTCCCACTACTGCTGTGACCAAGC	5100
QY	1455	LysLeuLeuThrSerThrThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr	1474
Db	5101	AAACTCCCTGACTTCTACCAACAAGTACTGCTTACCAACCAATTTACACTAGAACCA	5160
QY	1475	ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr	1494
Db	5161	GTTGCTTTGCCAGATTACACCAAGTGTGCACCTGGGCAAGTTTCTACCCCACTAGGACT	5220
QY	1495	ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro	1514
Db	5221	ACTACATCAGAGAGTGAACCTCGGAATGCTCTCCCTCCAGGCCACTCTAAGTAAAGCTCG	5280
QY	1515	ValLeuProValGlyThrGluLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro	1534
Db	5281	GTCGTGCGAGGGGAGTACTGAATCTTCAGCAAGTACTCTTACCAGGAGACACTGCCACT	5340
QY	1535	PheProGlyProSerSerLeuThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeu	1554
Db	5341	TTTCCAGGACCTTCTCTAACCCAGTCCAGCAACCTCTAAGGAGATCTTGAATGCTCAATGG	5400
QY	1555	ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValaGlyProValSerMet	1574
Db	5401	AGAGAGACACTTAATGCCAGAGATATATCAAGTACTTCTGGGCTTTGGCTGTGTCAATG	5460
QY	1575	AlaAlaProThrAlaAlleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnVal	1594

Dd	5461	CGCGCTCCAAAGCAATACAGAAAGCAAGAAACAGACCTTCAGAAAGGCTTTCACAGTC	5520
Qy	1595	LYSGIUGIYPRoValLEuAlaThrSerSerGIYlaGIYAlaPheLYsMeGIYARpHe	1614
Dd	5521	AAAGAGGCGCTGCTCCYAGCAACTAGTTTCAGAGACTGGTGTGTTTAAAGATGGAGCAATTT	5580
Qy	1615	GLINAlSerValAlaAlaASPGLYAlaGLInLYSGIUGIYLYsAsnLYsSerGIYAspAla	1634
Dd	5561	CAGGTTTCTGTGGACGACAGACGGGCCAGAAAGGGTATAAATATAGTCAGAAAGATGCA	5640
Qy	1635	LYsSerValHisPheGIJSerSerThrSerGIJSerSerValLEuSerSerSerSerPro	1654
Dd	5641	AAAGTCGTTCATTTTGAATCCAGACACCTCAGAGTCCTCAGTGTCTATCAAGTAAATAGTCC	5700
Qy	1655	GIJSerThrLEuValLYsPRoGIJNPRoAsnGIYleThrIlePRoGIYIleSerSerAsp	1674
Dd	5701	GAGAGTACTTGTGGAAACAGACCCGAAATGGCAATACATCCCTGGATTCCTTCAGAT	5760
Qy	1675	ValPRoGIJSerAlaHisLYsThrThralaSerGIYAlaLYsSerSerThrGIJNPRo	1694
Dd	5761	GTCGCAAGAGATGGCCCAAAACCTACAGCCTCAGAGGCAAAGTCAGACACTGGGCACACT	5820
Qy	1695	ThrLYsValGIYARpPheGIJNValThrThrThralaAsnLYsValGIYARpPheSerVal	1714
Dd	5821	ACCAAGGTGGACCTTTTCAGTGCACAACTACACAAACAAAGGGGCTGTTCTCTGTA	5880
Qy	1715	SerLYsThrGIJAspLYsIleThrAspThrLYsLYsGIJUGIYPRoValAlaSerPRo	1734
Dd	5881	TCAAAATCTAGGACCAAGATCATCTGACACAAAGAAAGAGACCAAGTGCATCTCTCCT	5940
Qy	1735	PheMeCAspLEuGIJNAlaValLEuPRoAlaValIlePRoLYsLYsGIJLYsPRoGIJN	1754
Dd	5941	TTTATGGATTGGAAACAGCTGTTCTTCCTGCTGTATACCAAGAAAGAGAACCTCGAA	6000
Qy	1755	LEuSerGIJNPRoSerHisLEuAsnGIYPRoSerSerAspPRoGIYAlaAlaPheLEuSer	1774
Dd	6001	CTGTGACAGCTTTCACATCTAAATGGCGCTTCTTGACCCGGAGGCGGCTTTTAAAGT	6060
Qy	1775	ARgAspValAspAspGIYSerGIYSerPRoHisSerPRoHisGIJNLEuSerSerLYsSer	1794
Dd	6061	AGGAGTGGATGATGGTTCGGTAGTCCACACTGCCCCATCAGCTGAGCTCAAGAGGC	6120
Qy	1795	LEuPRoSerGIJNAsnLEuSerGIJNLEuSerLysSerLysSerPheAsnSerSerTYMeSer	1814
Dd	6121	CTTCTCGACCAAGATCTAAAGTCAAAAGCCTTAAGTATTCATTAATCTCTCTTAATAGATG	6180
Qy	1815	SerAspAsnGIJSerAspIleGLIAspGIYAspLEuLYsLEuGIJNLEuARgARGLeuARg	1834
Dd	6181	AGGCACATAGTGCAGATGTGAAGATGAAGACCTTAAGTTAGAGCTGCCACGACTGCA	6240
Qy	1835	AspLYsHisLEuLYsGIJLleGIJNAspLEuGIJNLEuSerLysGIJNLYsHisLEuIleGIJNLEuSer	1854
Dd	6241	GATAAACATCTCAAAAGAGATTCAGAGACTGACAGTGCAGAGACATGAATTAATTAATCT	6300
Qy	1855	LEuYrThrLYsLEuGIYLYsValPRoPRoAlaValIleIlePRoPRoAlaAlaPRoLEu	1874
Dd	6301	TTGTATACCAACTGGGCAAGTGCCTCCCTGCTGTATATTCCCCCAGCTGCTCCCTT	6360
Qy	1875	SerGIYARgARgARgPRoThrLYsSerLYsGIYSerLYsSerSerARgSerSerSer	1894
Dd	6361	TCAGAGGAGAGACACACACCACTTAAGAAAGCAAGGCAAAATCTAATGCAAGACTTCC	6420
Qy	1895	LEuGIYAsnLYsSerPRoGIJNLEuSerGIYAsnLEuSerGIYGIJNLEuAlaAlaSerVal	1914
Dd	6421	TTGGGGATATAAACCCCCAGCTTTCAGGTACCTGTCTGTGTACAGGTGACGCTTCAGTC	6480
Qy	1915	LEuHisPRoGIJNThrLEuHisPRoPRoGIYAsnIlePRoGIJNLEuSerGIYGIJNAsnGIJN	1934
Dd	6481	TTTGACACCCGACGACACCTCTCAACCTCTGTGGACATCCAGAGTCCGGGCAAGATAG	6540
Qy	1935	LEuLEuGIJNPRoLEuLYsPRoSerPRoSerSerAspAsnLEuLYrSerAlaPheThrSer	1954

Db 6541 CTGTACAGCCCTTAAAGCATCTCCCTCCAGTACAGCACTCTATTACAGCTTCACCACT 6600

Qy 1955 ASPG1YAL1LESerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974

Db 6601 GATGTGTCCATTTCAGTACCAAGCCTTTCTGCTCCAGGTCAAGGAACACAGACACAAAC 6660

Qy 1975 ThrValG1YAlaThrValaenSerGlnAlaGlnAlaGlnProProAlaMetThrSer 1994

Db 6661 ACTGTGTGGGCAACAGTGAAGCAGCAAGCCGCCCAAGCTCAGCTCTGCAATGACGTCC 6720

Qy 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnTPAlaAsp 2014

Db 6721 AGCAGGAAGGACCATTCACAGATGACTTCCACAAAGTTGGTAGCAATTGGCCCGAGAT 6780

Qy 2015 AlMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrGlyProGly 2034

Db 6781 GCCATGATCTTCTCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6840

Qy 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysL1SerMetThrSerAsnLeuGly 2054

Db 6841 ATGGCAAGAAAGTTCTCTGACCTGGGCACTGTGATCTCAATGACCTCGAACTGGGT 6900

Qy 2055 GlySerAlaProL1SerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074

Db 6901 GGCTCTGCCCCCATCTCTGACATGACATCACTCTGAGTCACTTCCACCAAGTCTATG 6960

Qy 2075 CysProProGlnGlnTyrGlyPheProAlaThrProPheGlyAlaGlnTyrSerGlyThr 2094

Db 6961 TGCCCCCAGCAGCATGTGCTTCCAGCTCCATCCCATTTGGCGCTCAATGAGAGGAGAG 7020

Qy 2095 GlyLysProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114

Db 7021 GGTGGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7080

Qy 2115 AsnPheAsnL1SerAsnLeuGlnLysSerL1SerAsnProProGlySerAsnLeuArg 2134

Db 7081 AATTTCACATCAGCAATTTGACAGAAATCCATCAGCAACCCCGAGGCTCCAACTGCG 7140

Qy 2135 ThrThr 2136

Db 7141 ACCACT 7146

RESULT 2

PCT-US04-27459-1

; Sequence 1, Application PC/TUS0427459

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: PRKRNKS AS MODIFIERS OF THE RAC PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX04-060C-PC

; CURRENT APPLICATION NUMBER: PCT/US04/27459

; PRIOR FILING DATE: 2004-08-30

; PRIOR APPLICATION NUMBER: US 60/495,193

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 7149

; TYPE: DNA

; ORGANISM: Homo sapiens

PCT-US04-27459-1

Alignment Scores:

Pred. No.:	0	Length:	7149
Score:	10677.00	Matches:	2135
Percent Similarity:	89.63%	Conservative:	0
Best Local Similarity:	89.63%	Mismatches:	1
Query Match:	98.75%	Indels:	246
DB:	1	Gaps:	1

US-10-010-720-14 (1-2136) x PCT-US04-27459-1 (1-7149)

Qy 1 MetSerGlyGlyAlaAlaGlyLysGlnSerThrProGlySerLeuPheLeuSerPro 20

Db 1 ATGTCTGGGGGCGCCGCAAGAAAGCAGACAGACATCCCGGTTCCCTGTTCTCTGCGCG 60

Qy 21 ProAlaProAlaProLysAsnGlySerSerSerSerSerSerValGlyLysLeuGly 40

Db 61 CCGGCTCTGCCCCCAAGATGAGCTCAGGTCCGATTCCTCGGTGGGGGAGAAATCGGGA 120

Qy 41 AlaAlaAlaAlaAspAlaValThrGlyArgThrGlnGlyLysArgArgArgArgArgThr 60

Db 121 GCCGCGCGCGCGCACCTGTCGACCGGACAGAGACGAGAGATGACGCGCCCGCACACT 180

Qy 61 MetAspLysAspSerArgGlyAlaAlaAlaAlaThrThrThrThrThrThrThrThrThr 80

Db 181 ATGACAAAGAGACCGCGTGGGGCGCGCGGACCACTACCACTGACACCGCTTCTTC 240

Qy 81 ArgArgSerValL1LeuCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100

Db 241 CGCGGAGAGCTATCTGCGATCCATGACACAGCTGAGAGCTTCCGCGCTTCTCTCT 300

Qy 101 SerLeuProGlnProSerL1LeuProAlaAlaValProGlnSerAlaProProGlyProHis 120

Db 301 TCCCTGCCCCAGCCAGCATCCCGGCGCTGTCCGAGAGTCTCCACCGAGCCCGCAC 360

Qy 121 ArgGlnGlnThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140

Db 361 CGGAAAGAGCCGTGACCGCACCGCACCTTCCAGAGTACCCAGACCTTCCAGCCGCT 420

Qy 421 GCCGCCCCCTGGGGAAGACGCGTGGGGCGCTGCGCCCTCGACTGTGCCAGAGTACC 480

Qy 141 AlaAlaProGlyGlnGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160

Db 481 AGCAGAAAGACCGCCAGATGTCAGCCTTGTGGGGGCAAAAGAGGAGCGCCGCG 540

Qy 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGlnProProPro 180

Db 541 GCGAAGATGACAGCGCGCGGAGCGGCAAGAGGACAGAGAGAGAGAGAGAGAGAGAG 600

Qy 181 AlaArgSerLysSerGlyGlyLysSerAlaLysGlnProGlnGlnGlnGlnGlnGlnGln 200

Db 601 CAGGATATATCCAAAGAGCTGAGAGCCAGCCGTGGGAATGCTTAAAGATGCGCTT 660

Qy 201 GlnAspAspL1LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220

Db 661 CTCAAGTTGACATCGAAATCGCAGAGCTCTTTTAAAGCGGTCTCAAAAGGCTCTGAC 720

Qy 221 LeuLysPheAspL1LeuGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240

Db 721 ACTGAAACCAACCGTGAAGTGGCTGTGTGAATGCAAGATGCAAAATTTACAAAGTCT 780

Qy 241 ThrGlnThrThrValGlnValAlaAlaTyrCysGlnGlnGlnGlnGlnGlnGlnGlnGln 260

Db 781 GAGAGGACAGATTTAAAGAAAGCTGAATGTAAAGTCTTTCAGCATCCCAATATT 840

Qy 261 GlnArgGlnArgPheLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280

Db 841 GTTAGATTATGATTCCTGGAATCCACAGTAAAGAAAGAGTCAATGTTTGTG 900

Qy 281 ValArgPheTyrAspSerTyrGlnSerThrValLysGlyLysLysCysL1LeuVal 300

Db 901 ACTGAACTTATGAGCTGGAACCTTAAACGATCTGAAGAGCTTTAAAGTATGAG 960

Qy 301 ThrGlnLeuMetThrSerGlyThrLeuLysThrTyrLeuLysAspPheLysValMetLys 320

Db 961 ATCAAGTTCTAAAGCTGGTGGCGTCAATCTTAAAGTCTTCAAGTTCTTCAATCT 1020

Qy 321 L1LeuValLeuArgSerTyrCysArgGlnL1LeuLysGlyLeuGlnPheLeuHisThr 340

Db 1021 CGAAGTCACTATCATTCACCGGATCTTAAATGATGACAAATCTTATACCGGCGCT 1080

Qy 341 ArgThrProProL1LeuHisArgAspLeuLysCysAspAsnL1LeuPheLysThrGlyPro 360

Db 1081 ACTGCTCAGTCAAGATTTGAGACCTCGTCTGGCAACCTGAAAGCGGCTTCTTTTGC 1140

Qy 361 ThrGlySerValLysL1LeuGlyAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380

Db 1140 ACTGCTCAGTCAAGATTTGAGACCTCGTCTGGCAACCTGAAAGCGGCTTCTTTTGC 1140

```

QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
Db 1141 AAGAGTGTGATGATGATACCCAGAGTTCATGGCCCTGAGATGTATGAGAGAGAAATATGAT 1200
QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Db 1201 GAATCCGTTGACGTTATGCTTGGAGTGTGATGCTTGAGATGGCTATCTGAATAT 1260
QY 421 ProTyrSerGluCysGluAsnAlaIleGlnIleTyrArgValThrSerGlyValLys 440
Db 1261 CTTTACTGGAGTGGCCAAATGCTGCGCATCTACCGTCCGGTGGACGAGGGGTGAG 1320
QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleGluGlyCyste 460
Db 1321 CCAGCCAGTTTGGCAAGATGACATTCCTGAGTGAAGAAATATTATGAGATGCTA 1380
QY 461 ArgGluAsnLysAspGluArgTyrSerIleLysAspLeuAsnHisAlaPhePheGln 480
Db 1381 CGACAAAACAAAGATGAAAGATATTCATCAAGACCTTTTGAACCATGCTCTCCAA 1440
QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAla 500
Db 1441 GAGAAACAGAGATACGGGTGAAATTAGACAGAGAAATGATGAGAAATAATAGCCATA 1500
QY 501 LysLeuTyrPheAspValIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsn 520
Db 1501 AAATTATGCTACGATATGAAATATTAAAGAAATTTAAAGGAAATATCAAAAGATATGAA 1560
QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGlu 540
Db 1561 GCTATTAGTTTCTTTTGAATTTAGAGAGATGCTCCAGAAAGATGTTGCAACAAAGAAATG 1620
QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 1621 GTTAGAGTCTGGATGTCTGTGAAAGTGTATACAAAGCATGCTAAAGCTATCAAGAC 1680
QY 561 ArgValSerLeuIleLysArgGlyArgGluGlnArgGluLeuValArgGluGluGln 580
Db 1681 AGAGTATCATTTAATTAGAGAGAAACGAGACGCGCATGCTGATCGGAGAGACAAAGA 1740
QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnValGluGlnSerSerAlaSerGln 600
Db 1741 AAAAAAAGAGAGAGAGACAGCTCTAAACAGCAGGTAGAACATCCAGTCTCCAG 1800
QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 1801 ACGAGATCAAGCAGCTCCCTTCTGCTAGACCGGCACTACTGCTTCTACCACTTCA 1860
QY 621 AlaSerValSerThrGluValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Db 1861 GCTTCAGTTTCTACACAGTAGAACCTGAAGAACTGAGGCACTCAACATCAACAACTA 1920
QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 1921 CAGTACCAAGACCACTATATCTGTATCTGATGGACGTTGACAGTGTGACAGGGA 1980
QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 1981 TCTCTGTCTTCAACAGATCTCAGTAGACAGCCACACAGATTCATATGGTTCCAA 2040
QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2041 CATGAACAGCACATCTTACAGACAGCTCCAGGGGATATACCTTCTACTGTCACAGCA 2100
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 2101 CAGTCTCAGCCCATGGGATATATCAACCTCAAGGTGGCACAGGGGAGACCGAGGT 2160
QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 2161 CAGCATCTCTCAAGTACCTTAACAGGGGTTCTATCTTCCCAACCATCAACATCTCTCAG 2220

```

```

QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 2221 CAGCAGCAGGAAATACAGACAGCCCTCTCAACAGACAGTCAATTCATCTTCA 2280
QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 2281 CAGCATCAACCTCCAGTGAAGGCCACTACTGACAGCCAGTGAAGTCAAGCTTCAAGCTCCA 2340
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 2341 CAACTTGGCTCAAGTATCAGCTGSAACAGCTTCCAGTTTCCACAGCTACCAACT 2400
QY 791 ----- 791
Db 2401 ATCAAGCCGACCTCAGATCCAGTTCGACACACAACTCGTTGTTCCAGTCCACT 2460
QY 791 ----- 791
Db 2461 GGTGCTCATTTCTTCCAGTGGAGACAGCGCTCCCTACTCCCTGCTCTCAGTACCT 2520
QY 791 ----- 791
Db 2521 GTCTCTAGATTCCTCATATCAACTCTCATGTGTCTACGCTCAGACAGTTCCTCATCC 2580
QY 791 ----- 791
Db 2581 CTTCCTCATCATATGACGACGCTGGCATTACTAGCCTCTGCTCAGCTTGCTTCATCTGCT 2640
QY 791 ----- 791
Db 2641 ACAACAGCTCGATCCCGGGGTATCAACTGTGTCTCTAGTACGTTCCAACTCTCTG 2700
QY 791 ----- 791
Db 2701 CAGCTGTGACTGACTGCTGCCAAGTCAAGTTTACCCACAGCTCTTCAACAGCAGATTCA 2760
QY 791 ----- 791
Db 2761 TCATGGGAATACCACTGTAACCTTGGACAAAGCTGTGAGTTCCATCTTCTCTGAGAT 2820
QY 791 ----- 791
Db 2821 GTTCTGTACAGGGCTTCCCACTCGACTGCAACACAGTACCCAGAGATTCAAATATT 2880
QY 791 ----- 791
Db 2881 GCTCCCTCTTCAACAGTGGCTTCTGTTGATCATCATCTTCAAGTCTTATCCCTCATG 2940
QY 791 ----- 791
Db 2941 CCGACAGAACTACTGGCTACCTGGGTACTTTCACAGTGTGACGCTTATGTGAA 3000
QY 791 ----- 791
Db 3001 TCAATCTTTTACTTCTCATAGGTGTGTAGAGAGACAGGTTCAAGTCCAGCCAGCA 3060
QY 792 ----- 792
Db 3061 GGGAGTTTACACAGCCCCCACTACATCTCCACAGACAGTTTGGAGAGTACTCAG 3120
QY 795 GlyValSerGlnAlaIleProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
Db 3121 GAGTCTCTCAGTTCTCTCTGACAGGCCAGTTGACAGTACAGACAGCCCAAGCTTACCCAG 3180
QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
Db 3181 CCAGCCACTTGTGCTTCTCTGTAGACAGTGCATTCAGATTTCTTCAAGTATGAGT 3240
QY 835 AspGlyAsnGluAsnValProSerSerSerArgHisGlnGlyArgThrThrLysArg 854
Db 3241 GATGGCAATGAGAACCTCCATCTTCCAGTGAAGGCAATGAAAGAACTTCAAAAACGG 3300
QY 855 HisTyrArgLysSerValArgSerArgSerArgHisGluLysThrSerArgProLysLeu 874

```

Db	3301	CATTACCGAAATCTGTAGAGAGCGCTTCGACATGAAATAAACTTCACCCCAAAATTA	3360	Db	4381	GGGGGGAGTACTGCTACCCCGAGTCTTAAGCTTCAGCTGTATCTCAGCAGGACGA	4440
Qy	875	Arg11IleuAsnValSerAsnLysGlyAspArgValValGluCysGlnIleuGluThrHis	894	Qy	1235	GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrSerPheProSer	1254
Db	3361	AGAAATTTGAAATGTTTAAATTAAGAGACCGAGTAGTAAATGTCATTAATGAGACTCAT	3420	Db	4441	GGCAGACACTACTGTGGAGGCACTTAACATCACTGTTTTCACCACTTATTCACCAAGC	4500
Qy	895	AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluIleAla	914	Qy	1255	ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrThrProThrLeuAlaGlu	1274
Db	3421	AATAGGAAATGGTTACATTCATTAATTTGACCTAGATGTCACAACCCCGAGAGATAGCA	3480	Db	4501	ACAGCTTCACAGCTGATCCATTCAGCTTAGCAGCAGATACCTCTACTCTTACCTTAAGTGAA	4560
Qy	915	ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgGluSerPheValAspGln	934	Qy	1275	ThrValValAlaSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu	1294
Db	3481	ACAAATTGCTGAACAATGACTTATCTAGCAATAGAGAGAGAGTGGTTGTGATCAAA	3540	Db	4561	ACCGGTGTAGTGGCCACACTCCTACATGATTAAGACATCTCATGACGTACACTGGAATG	4620
Qy	935	ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValGluPro	954	Qy	1295	AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyr	1314
Db	3541	GTGCGAAGAAATTTATTGAAAAAGCTGATGAATGCTCAGTAGAGATGTCAGTGGAAACA	3600	Db	4621	GGTTTCTCCTCTCTGACCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4680
Qy	955	GluGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer	974	Qy	1315	IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro	1334
Db	3601	GAGGGTATACAGGATTTGAGAGAGTCTACAGAAAGATGACTATGCTTTTCAAGTTCT	3660	Db	4681	ATTCTCAGCTGTGGGGCTGCATCTTGTGTCATTCCATCAATGATAGCTTCTACTCTCT	4740
Qy	975	GlnLysLeuGluGlyGlyPheLysGlnProIleProAlaSerSerMetProGlnIle	994	Qy	1335	IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle	1354
Db	3661	CAGAAATTTGGAGAGAGGTTCAAAACAACCAATTCCTGCTTCCATGCCACAGCAAAATA	3720	Db	4741	ATTCTTCCCAACAGCAGACGACCTTCTTACACTTTATTAACCCCAAGTACTGATATC	4800
Qy	995	GlyIleProThrSerSerLeuThrGlnValHisSerAlaGlyArgGlyPheIleVal	1014	Qy	1355	ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer	1374
Db	3721	GGCAATCTTACCGATTTTAACTCAAGTGTTCATTTCTCGGAGAGCGGTTTATATG	3780	Db	4801	CCACCTTGGTATACGCTGTGGCCAAATGTCCTGCTGTACAGACAGACATTAATTCATAGT	4860
Qy	1015	SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluIleThrAsp	1034	Qy	1375	GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer	1394
Db	3781	AGTCTGTGCGCAGAAAGCCGATTAACGAAATCAAAAGTTTCCCACTGAAATACGAT	3840	Db	4861	CACCTCAACACAGCTTGTGCTTCCCAACAGCCCAATACCTATGTCTCTGAATGATTTCT	4920
Qy	1035	ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer	1054	Qy	1395	AspThrGlnProLysAlaProGlyIleAspAspIleLysThrLeuGluGlyLysLeuArg	1414
Db	3841	ACAGTTGCTCCTCTTACAGCTCAGAGCCCTGGAATGAACTGTCTCACTCGATCATCTCC	3900	Db	4921	GATACCAACCCCAAGCTCTGGAATGAGATGAATGAAGCTGTAGAGAAAGCTGCGG	4980
Qy	1055	LeuSerLeuGlnGlnAlaPheSerGluLeuArgAlaGlnMetThrGluGlyProAsn	1074	Qy	1415	SerLeuPheSerGlnHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThr	1434
Db	3901	CTTAAGTCTACACAGGCGCTTTCTGAACTTAGAGAGTCCCAATAGACAGAAAGGCCAAC	3960	Db	4981	TCTCTGTTCAGTAAACAGAGCTCATCTGAGCTCAGATCTCTCTGCTTCACTGAGAAC	5040
Qy	1075	ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu	1094	Qy	1435	SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer	1454
Db	3961	ACAGCACTCCAAACTTACTGATACAGAGCAACATTTCCAGTAGTAACTCTTCTTCTTA	4020	Db	5041	TCACTAGTCAATAGAGGACTGTACACACAGGATCCCACTACTGCTGTGACCAAGC	5100
Qy	1095	SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer	1114	Qy	1455	LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr	1474
Db	4021	AGTAGCATTTGCTGAGTCCCAACACAGCAGCAGCCACACACAGTCCCTGCAACAAAGC	4080	Db	5101	AAACTCTGACTTCTTACCAAGTACTTGTCTTACCAACCAATTTACCACTAGGAGACA	5160
Qy	1115	SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu	1134	Qy	1475	ValAlaLeuProValThrProValAlaThrProGlyGlnValSerThrProValSerThr	1494
Db	4081	AGCCCTCTAATGACATTTCCACATCGTAATTCAGTGAAGTTTAAAGGCCACCTGAA	4140	Db	5161	GTTCCTTTGCCAGTTTACACAGTGTGCACCTGGGCGAATTTCTTACCCAGTCCAGTACT	5220
Qy	1135	GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLysLeuProIle	1154	Qy	1495	ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro	1514
Db	4141	GAGGGATTTGCTGAGTGGCCACAGCAGCAGGTGTGTAATCTCAGGTGCTCCCATTA	4200	Db	5221	ACTTACATCAGAGATGAAGAACTGGAATGCTCCCTCCACAGCACCTGTAATTAAGGCTCG	5280
Qy	1155	ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla	1174	Qy	1515	ValLeuProValAlaGlyThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPro	1534
Db	4201	CCACCTGTGCTGTAATCACAGTACTTCCAGCTAGTTCATCAAGTATCACCAATCTCGCA	4260	Db	5281	GTGCTGCAGTGGGTACTGAATCTTCAGCAGGTACTTACCCAGCAGGAGCTGCCACT	5340
Qy	1175	ValValSerIleSerThrThrThrSerProSerLeuGlnValProThrSerThrSerGluIle	1194	Qy	1535	PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnLysLeuAspAlaGlnLeu	1554
Db	4261	GTGTGCTCAATATCTACATATCCCGTCACTTCAATCCCAATCCCAATCCCAATCCGAATC	4320	Db	5341	TTTCCAGGACCTTCTCTTAACTCCAGTCCCAAGACCTTACAGAGATCTTGAATGCTCAATG	5400
Qy	1195	ValValSerSerThrAlaLeuLysProSerValThrValSerAlaThrSerAlaSerAla	1214	Qy	1555	ArgArgThrLeuSerProGluLysIleThrValThrSerAlaValAlaGlyProValSerMet	1574
Db	4321	GTGTGTTCTGATAGACACTGTATCTTGTAGTAAATTTTCAAGAACTTCAAGCTCTCGCA	4380	Db	5401	AGAAAGAACCTTATGTCAGAGATTAATCAAGTACTTCTGCGTGGTCTCTGTGTCATG	5460
Qy	1215	GlyGlySerThrAlaThrProGlyProLysProProAlaValValSerGlnAlaAla	1234	Qy	1575	AlaAlaProThrThrAlaIleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnVal	1594
Db				Db	5461	GGCGCTCCACAGCAATCAACAGAGAGCAAGAAACAGCTCTGAGAGGCTGTTCTCAAGTCT	5520

QY 1595 LysGluGlyProValLeuAlaThrSerSerGlyValGlyValPheLysMetGlyArgPhe 1614
DB 5521 AAGAGAGCCCTGCTCAGCACTAGTTCAGAGCTGGTGTGTTTAAAGAGGCGATTT 5560
QY 1615 GluValSerValAlaAlaAspGlyValGluGlyLysAsnLysSerGluAspAla 1634
DB 5581 CAGGTTCTGTGAGAGAGAGGTCGCCAGAAAGGGTAAATTAAGTCAGAGATGCA 5640
QY 1635 LysSerValHisPheGluSerSerThrSerGluSerValLeuSerSerSerPro 1654
DB 5641 AAGCTGTTCATTGTGAATCCAGACCTCAGAGCTCAGCTCATCAAGTAGAGTCCA 5700
QY 1655 GluSerThrLeuValLysProGluProAsnGlyLThrLysPheGlyLysSerAsp 1674
DB 5701 GAGAGTCTTGTGGAAACCAAGCCCAATGCAATACCTCCTGGTATCTTTCAGAT 5760
QY 1675 ValProGluSerAlaHisLysThrThrAlaSerGluAlaLysSerAspThrGlyGlnPro 1694
DB 5761 GTGCCAGAGAGTGCCCAAAACTAGTCCCTCAGAGGCAAGTCAGACACTGGGCACT 5820
QY 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
DB 5821 ACCAAGGTGACGCTTTCAGGTGACCACTACAGCAAAAGGGGTGCTTCTCTGTA 5880
QY 1715 SerLysThrGluAspLysLysLysThrAspThrLysLysGlyProValAlaSerProPro 1734
DB 5881 TCAAAAACAGAGCAAGATCACTGACACAAAGAAAGAGCAAGTGGCATCTCTCT 5940
QY 1735 PheMetAspLeuGluGlnAlaValLeuProAlaValLysProLysGlyLysProGlu 1754
DB 5941 TTATGATTTGGAAACAGGCTTCTCTGCTGTGATACCAAGAAAGAAAGCACTGAA 6000
QY 1755 LeuSerGluProSerHisLysAsnGlyProSerSerAspProGluAlaAlaPheLeuSer 1774
DB 6001 CTGCAAGGCTTACACTTAATGAGGCGCTTCTTCAACCCGAGGCGCTTTTAAAGT 6060
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
DB 6061 AGGAGTGTGATGATGTTCCGGATGTCACACTGCCCATCAGCTGAGCTCAAAAGC 6120
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerLysMetSer 1814
DB 6121 CTTCCTCAGCAAGATCAATCAAGTCAAAAGCCTTAGTAATCATTTAACTCTTACATAGT 6180
QY 1815 SerAspAsnGluSerAspLysLysLysGluAspGluAspLeuLysLeuAlaGluLeuArg 1834
DB 6181 AGCGACATAGATGATGATGATGAAAGATGAAAGATTAAGTGAAGCTGCGACGACTACGA 6240
QY 1835 AspLysHisLeuLysGluLysLysGlnAspLeuSerArgGlnLysHisGluLysSer 1854
DB 6241 GATTAACATCTCAAAAGATTCAGAGACTTCAGAGTCCAGAGTCCAGAAATTTGATCT 6300
QY 1855 LeuLysThrLysLeuGlyLysValProProAlaValLysLeuProAlaAlaProLeu 1874
DB 6301 TTGTATATCCAAACTGGGCAAGTGCCCTGCTGTATTTATCCCCAGCTGCTCCCTT 6360
QY 1875 SerGlyArgArgArgArgProThrLysSerLysGlySerLysSerArgSerSerSer 1894
DB 6361 TCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6420
QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
DB 6421 TTGGGGAGATAAAGCCCAAGCTTTCAGGTACCTGCTGCTGAGATGACGCTTCACTG 6480
QY 1915 LeuHisProGlnGlnThrLysHisProProGlyAsnLysProGluSerGlyGlnAsnGln 1934
DB 6481 TTTCACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6540
QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
DB 6541 CTGTTCAGAGCCCTTAAAGCATCTCTCCAGTGAACCACTTATTCAGGCTTCAACAGT 6600

QY 1955 AspGlyValAlaLysSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
DB 6601 GATGTCTCATTTCAGTACCAAGCTTCTTCTGCTCAGGTTCAGAGAACAGAGCAACAAC 6660
QY 1975 ThrValGlyValAlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSer 1994
DB 6661 ACTGTGGGCAACAGTGAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 6720
QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnThrPalaArgAsp 2014
DB 6721 AGCAGAGAGGAGCAATTCACAGATGACTTCCACAAAGTTGGTAGACAAATGGGCGCGAGAT 6780
QY 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnArgGlyProGly 2034
DB 6781 GCGATGATCTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6840
QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysLysSerMetThrSerAsnLeuGly 2054
DB 6841 ATGGCAAGAGAGATCTCTGCACTGGGCAACTGTCATCTCCAGACTCGAAGCTGGGT 6900
QY 2055 GlySerAlaProLysSerAlaAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074
DB 6901 GGCTGTGCGCCCATCTCTGAGCATGACTTCTGATGATCACTTCCACCAAGTCTATG 6960
QY 2075 CysProProGlnGlnArgLysGlyPheProAlaThrProPheGlyAlaGlnThrSerGlyThr 2094
DB 6961 TGCCCCCAAGAGAGATGAGCTTTCAGCTAACCCATTTGGCGCTCATATGAGTGGGAGC 7020
QY 2095 GlyGlyProAlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
DB 7021 GTGGGCCAGACACACAGCACTGGCCAGTTCACACTGTGGGAAGTGCCTCTTGCAG 7080
QY 2115 AsnPheAsnLysSerLeuGlnLysSerLysSerAsnProProGlySerAsnLeuArg 2134
DB 7081 AATTTCAACATCGACATTTGGCAAAATCCATGACCAACCCCGAGCTTCAACTGGCG 7140
QY 2135 ThrThr 2136
DB 7141 ACCACT 7146

RESULT 3
US-10-491-467-54
Sequence 54, Application US/10491467
GENERAL INFORMATION:
APPLICANT: INCYTE CORPORATION; BANDMAN, Olga;
APPLICANT: BAUGHN, Mariah R.; BECHT, Shanya D.;
APPLICANT: BOROWSKY, Mark L.; DUGGAN, Brendan M.;
APPLICANT: EMERLING, Brooke M.; FORSYTHE, Ian J.;
APPLICANT: GANDHI, Ameena R.; GORVAD, Ann E.;
APPLICANT: GRIFPIN, Jennifer A.; GURURAJAN, Rajagopal;
APPLICANT: HAFALIA, April J.A.; KHAN, Farrah A.;
APPLICANT: LAL, Preeti G.; LEE, Ernestine A.;
APPLICANT: LEE, Soo Yeun; LINDQUIST, Erika A.;
APPLICANT: LU, Dying Anna M.; LU, Yan;
APPLICANT: MARQUIS, Joseph P.; NGUYEN, Daniel B.;
APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi;
APPLICANT: RECIPION, Shirley A.; RICHARDSON, Thomas W.;
APPLICANT: SWARNAKAR, Anita; TANG, Y. Tom;
APPLICANT: THORNTON, Michael B.; TRAN, Uyen K.;
APPLICANT: CHAWLA, Narinder K.; WARREN, Bridget A.;
APPLICANT: YANG, Junming; YAO, Monique G.;
APPLICANT: YUE, Henry; ZEBARUDIAN, Yeganeh
TITLE OF INVENTION: KINASES AND PHOSPHATASES
FILE REFERENCE: PF-1244 USN
CURRENT APPLICATION NUMBER: US/10/491,467
CURRENT FILING DATE: 2004-03-31
PRIOR APPLICATION NUMBER: PCT/US02/33723
PRIOR FILING DATE: 2002-10-17
PRIOR APPLICATION NUMBER: US 60/345,474
PRIOR FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: US 60/343,910
PRIOR FILING DATE: 2001-11-02
PRIOR APPLICATION NUMBER: US 60/333,098

Db 4237 CTGCATCTTGGTCATTCATCAGTAGTAGCTTCTAATCTCTTATCTTCCCAAGCAGCA 4296
Qy 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db 4297 GGACCTACTTCTACACCTTTATTAACCAAGTACTTATCCACCCCTGGTATACAGCT 4356
Qy 1361 ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 4357 GTTGCATATGCTGCTGTACAGACAGACTAATTCATAGTCAGCTCAACCAAGCTTGG 4416
Qy 1381 LeuProAsnGlnProHisThrHisCysProGlnValAspSerAspThrGlnProValAla 1400
Db 4417 CTTCACCAACGAGCCCACTACTATGCTCTGAAGTAACTTCTGATACCAACCAAGCT 4476
Qy 1401 ProGlyIleAspAspIleLeuThrLeuGlnGluLeuArgSerLeuPheSerGluHis 1420
Db 4477 CTTGGAATTCATATCAATTAAGACTCTAGAGAAAAGCTGCGGTCTCTGTTCAATGAACAC 4536
Qy 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGlnSer 1440
Db 4537 AGCTCATCTGGAGCTCAGACTGCTCTGTCTCACTGGAGACTCACTAGTCATAGAGAGC 4596
Qy 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerIleLeuThrSerThr 1460
Db 4597 ACTGTACACACGAGCATCCCACTACTGCTGTGACCAAGCAAACTCTGACTTCTTACC 4656
Qy 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
Db 4657 ACAAGTACTGTTTACCAACCAATTTACCTAGAGAACAGTTGCTTTGCCAGTTTACA 4716
Qy 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValLys 1500
Db 4717 CCAGTGTGCACACTGGGCAAGTTTCTACCCAGTCAGACTACTACATCAGAGAGTAAA 4776
Qy 1501 ProGlyThrAlaProSerIleProProLeuThrIleAlaProValLeuProValGlyThr 1520
Db 4777 CCTGGAACTGCTCCCTCCAGACCACTTACTAAGGCTCCGGTGTGCTGCCAGTGGGACT 4836
Qy 1521 GluLeuProAlaGlyThrLeuProSerGlnGlnLeuProProPheProGlyProSerLeu 1540
Db 4837 GAATCTTCAGACAGTACTTCAACCAAGCAGAGCTGCGACCTTTTCCAGAGACCTTCTCTA 4896
Qy 1541 ThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgThrLeuSerPro 1560
Db 4897 ACCCAGTCCAGCAACTCTAGAGGATCTTGAATGCTCAATTGAGAGAAACACTTATGCCA 4956
Qy 1561 Gln***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle 1580
Db 4957 GAGATTATCACAGTGACTTCTGGGGTGTCTGTGTCCATGGCGGCTCCAAACAGCATC 5016
Qy 1581 ThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGlnGlyProValLeu 1600
Db 5017 ACGAAGCAGGAACACAGCTTCAGAAAGGAGTCTTAAGTCAAAAGAAAGCCCTGTCTTA 5076
Qy 1601 AlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAlaAla 1620
Db 5077 GCAACTAGTTCAGGAGCTGTGTGTTTAAATGAGGAGATTTCAGGTTTCTGTTGCACGA 5136
Qy 1621 AspGlyValaGlnLysGlnGlyLysAsnLysSerGlnAspAlaLysSerValHisPheGln 1640
Db 5137 GACGGTCCCAAGAAAGGGGTAAATAATAGTCAGAAATCAAAAGTCTGTCTCAATTTTGA 5196
Qy 1641 SerSerThrSerGlnSerSerValLeuSerSerSerSerProGlnSerThrLeuValLys 1660
Db 5197 TCCAGACCTTCAGAGTCTCAGTGTCTTCAAGTATGTCACAGAGTACTTGGTATAA 5256
Qy 1661 ProGlnProAsnGlyIleThrIleProGlyIleSerSerAspValProGlnSerAlaHis 1680
Db 5257 CCAAGAGCGAATGGCATTAACCATCTCGTATCTCTTCAAGATGGCCAGAGAGGCCAC 5316
Qy 1681 LysThrThrAlaSerGlyAlaLysSerAspThrGlyGlnProThrLysValGlyArgPhe 1700
Db 5317 AAAACTACTGCTCTAGGGCAAGTCAAGACATGGGGAGCCCTTCAAGAGTTTGAAGTTT 5376

Qy 1701 GlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThrGlnAspLys 1720
Db 5377 CAGGTACAACTACACAAACAAAGGGGTGCTTCTGTATCAAAAACTGAGAGCAAG 5436
Qy 1721 IleThrAspThrLysLysGlnGlyProValAlaSerProProPheMetAspLeuGln 1740
Db 5437 ATCATTGACACAAAGAAAGAGGACAGTGGCATCTCTCTTTATGATTGATTTGAACAA 5496
Qy 1741 AlaValLeuProAlaValIleProLysLysGlnLysProGlnLeuSerGlnProSerHis 1760
Db 5497 GCTGTTCTTCCTGCTGTGATACCAAGAAAGAGAACCTGAATCTCAGAGCTTCAAT 5556
Qy 1761 LeuAsnGlyProSerSerAspProGlyAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db 5557 CTAAATGGGCGCTTCTGTACCCCGAGGCGCGCTTTTAAAGTGGAGTGTGATGATGT 5616
Qy 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
Db 5617 TCCGTTAGTTCACACTGCGCCCATCAGCTGAGCTCAAGAGCTTCTAGCCAGAACTTA 5676
Qy 1801 SerGlnSerLeuSerSerPheAsnSerSerTyrMetSerSerAspAsnGlnSerAsp 1820
Db 5677 AGTCAAGCTTATGTAATTCATTAACTCTCTTACATGATGAGCAATGATGATGAT 5736
Qy 1821 IleGlnAspGlnAspLeuLysLeuGlnLeuArgArgLeuArgAspLysHisLeuLysGln 1840
Db 5737 ATCGAAGATGAACACTTAAGTTAGAGCTGCCAGCTACAGATAAACATCTCAAGAG 5796
Qy 1841 IleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLeuThrLysLeuGly 1860
Db 5797 ATTCAGAGACTGACAGTGCAGAGCATGAATGAATCTTGTATACCAACTGGGC 5856
Qy 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyValArgArgArg 1880
Db 5857 AAGTGCCTCCCTGCTGTATATTTCCCACTGCTCCCTTTCAGAGAAACAGCA 5916
Qy 1881 ProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlnAsnLysSerPro 1900
Db 5917 CCACTTAAGGCAAGCAAGCAGCAATCTAGTCAGAGCATGTTCTGGGGATTAAGCCCC 5976
Qy 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
Db 5977 CAGCTTTCAGTAACTGTCTGTGTGTCAGAGTCCAGTTCACTTGCACCCCGACAGACC 6036
Qy 1921 LeuHisProProGlyLysLeuIleProGlnSerGlyGlnAsnGlnLeuGlnProLys 1940
Db 6037 CTCACCTTCCTGGCAACATCCAGAGTCCGGGCAAGTCACTGTACAGCCCCCTTAAG 6096
Qy 1941 ProSerProSerSerAspAsnLeuTyrSerAlaPheThrSerAspGlyAlaIleSerVal 1960
Db 6097 CCATCTCCCTCCAGTACAACTCTATTCAGCTTACCAAGTATGATGTCATTCAGTA 6156
Qy 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
Db 6157 CCAAGCTTTCCTGCTCAGGTCAAGGAAACAGACAGCAAAACCTGTTGGGCAACAGTG 6216
Qy 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
Db 6217 AACAGCAAGCCGCCCAAGCTCAGCTCTCCCATACATGTCACAGAGAAAGGCACTTC 6276
Qy 2001 ThrAspAspLeuHisLysLeuValAspAsnTyrAlaArgAspAlaMetAsnLeuSerGly 2020
Db 6277 ACGATGACTTGCACAAAGTTGTAGAACATTTGGGCCGAGATCCATGATCTCTCAGGC 6336
Qy 2021 ArgArgGlySerLysGlyHisMetAsnTyrGlnGlyProGlyMetAlaArgLysPheSer 2040
Db 6337 AGCAGAGGAAGCAAAAGGCAATGAATTAAGAGGCGCTGGAATGCCAAGAGATCTCT 6396
Qy 2041 AlaProGlyGlnLeuLysLysSerMetThrSerAsnLeuGlyGlyLysSerAlaProIleSer 2060
Db 6397 GCACTGGGCACTGTGATCTCCATGACCTCAACCTGGAGTGTGCTGCCCATCTCT 6456

QY 975 GlnIysLeuGlnGlyGluPheIysGlnProIleProAlaSerSerMetProGlnGlnIle 994
Db 4516 CAGAAATTCAGAGAGAGAGTTCACCAACCAATTCCTGCTTCCTCATTCACCAAGAAATA 4675
QY 995 GlyIleProThrSerSerLeuThrGlnValValHisSerAlaGlyArgArgPheIleVal 1014
Db 4676 GGCAATTCCTACAGATTCCTTAACTCAAGTTCATTCCTGCGGGAAGCGGTTTAAATGTC 4735
QY 1015 SerProValProGluSerArgLeuArgGluSerIysValPheProSerGluIleThrAsp 1034
Db 4736 AGTCTGTCGCGCAAGAGCCGATTCAGCAATCAAAAGTTTCCTCCAGTGAATATACAGAT 4795
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db 4796 ACAATGCTGCTCCTACAGCTCAGAGCCCTCGAATGAATCTGCTCCTCATCTGCATATCC 4855
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsn 1074
Db 4856 CTTAGCTTACAAAGAGGCTTTCTGAACCTTAGAGCTGCCAAATGACAGAAAGAACCCAAAN 4915
QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu 1094
Db 4916 ACAGACCTCCAAACTTATGATACACAGACCAACCAATTCAGATGATACCTCTTCTTA 4975
QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 4976 AGTAGCATTCGTGAGTCCCAACACACAGACAGACAGACAGACAGACAGACAGACAGAC 5035
QY 1115 SerProProAsnAlaPheSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
Db 5036 AGCCCTCTTAATGACATTCCTCAACATCAGTAAATTCAGCTGAGTTACAGTCCACACGAA 5095
QY 1135 GluGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyGlyLeuProIle 1154
Db 5096 GAGGGAGATTCGTGAGTCCACACAGACAGAGTGTGTAACTTACAGTGTGCTCCCATTA 5155
QY 1155 ProProValSerGluSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 5156 CCACCTGTGTGATACACAGTACTTCCAGCTGATTTCAAGTATCACAAATACCTGCA 5215
QY 1175 ValValSerIleSerThrThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
Db 5216 GTTGTCTCAATCTACTACATCCCGCTCACTCAAGTCCCACTCCACATCCAGATC 5275
QY 1195 ValValSerSerThrAlaLeuThrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 5276 GTTCTTTCTAGTACAGACCTGTATCTTCACTTCAAGTATTCAGACCTTCAAGCTCTGCA 5335
QY 1215 GlyIysSerThrAlaThrProGlyProIysProProAlaValValSerGlnGlnAla 1234
Db 5336 GGGGGAGTACTGCTACCCAGGCTCTAAGCTCCAGCTGATATCTCAGAGGAGCA 5395
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 5396 GGGAGCACTACTGGGAGCACTTAACATCACTTCTTACACACACTTCAATTCACAGC 5455
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlu 1274
Db 5456 ACAGCTTCACAGCTGTGSCATTCACCTTAGACAGACGATCTTCTACTCTTAACTTAACTGAA 5515
QY 1275 ThrValValValSerAlaHisSerLeuAspIysThrSerHisSerSerThrThrGlyLeu 1294
Db 5516 ACCGTGTAGTTCAGCAGACCTCACTAGATTAAGACATCTCATAGCAGTACCACTGATG 5575
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerThr 1314
Db 5576 GCTTTTCCTCTCTGACACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5635
QY 1315 IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 5636 ATTTCACAGCTGTGGGCTGCATCTTGTGTATTCATCATGATGATAGCTTCTACTCTCT 5695
QY 1335 IleLeuProGlnAlaAlaGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 5696 ATTTCCTCCAGAGAGAGGACTTACTTACACTTATATATCCCAAGTACCTGATATC 5755
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnThrLeuIleHisSer 1374
Db 5756 CCACCTTGTGATACAGCTGTTGCCAATGTGCTCTGTGACAGACAGACACTAATCATATGT 5815
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValAspSer 1394
Db 5816 CAGCTCAACCAAGCTTGTCTTCCCAACAGCCCATCTCATTTGTCTCGAAGTGAATTC 5875
QY 1395 AspThrGlnProIysValProGlyIleAspAspIleIysThrLeuGlnGluIysLeuArg 1414
Db 5876 GATACACACCCAAAGCTCTCGGATTTGATGATACATTAAGATCTTAGAAGAAAGACTGGG 5935
QY 1415 SerLeuPheSerGluHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThr 1434
Db 5936 TCTCTGTTCAGTGAACAGCTCATCTGAGCTCAGATGCTCTGTCTCACTGGAGACC 5995
QY 1435 SerLeuValIleGluSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 5996 TCACTAGTCAATAGAGACACTGTACACAGCAGCATCCCACTACTGCTGTGACACAGC 6055
QY 1455 LysLeuLeuThrSerThrThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 6056 AAATCTCGACTTCTACACAGATGATCTGTATACACCAACATTTTACCTAGAAACA 6115
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 6116 GTTGTCTTTCAGTTCACACAGAGTGTCAACCTGGGCAAGTTTCTACCCAGTACAGACT 6175
QY 1495 ThrThrSerGlyValIysProGlyThrAlaProSerIysProProLeuThrIysAlaPro 1514
Db 6176 ACTACATACAGATGAAACCTGGAATCTCTCTCCAGACCACTTCTTAACCTTAAGGCTTCG 6235
QY 1515 ValLeuProValGlyThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPro 1534
Db 6236 GTGTGTCAGTGGGTATGAACTTCAGAGATCTTACCCAGCAGAGCTGTCCACT 6295
QY 1535 PheProGlyProSerLeuThrGlnSerGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 6296 TTTCCAGAGACTTCTCTTACCCAGCTCCAGCAACCTTCAAGATCTTGATGTCTCAATTC 6355
QY 1555 ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValGlyProValSerMet 1574
Db 6356 AGAAGACACTTATGTCAGATGATATACAGTATCTTCTGCTGTGCTCTGTGCTCATG 6415
QY 1575 AlaAlaProThrAlaIleThrGluAlaGlyThrGlnProGlnIysGlyValSerGlnVal 1594
Db 6416 GCGGCTCYAACAGCAATCACRAAGAGAGCAACAGCTCAGAAAGGTGTTCACAGTC 6475
QY 1595 LysGluGlyProValLeuAlaThrSerSerGlyAlaGlyValPheIysMetGlyArgPhe 1614
Db 6476 AAGAAGGCCCTGTCTCTAGCACTAGTTCAGAGCTGTGTCTTTTAAAGTGGAGCGATTT 6535
QY 1615 GlValSerValAlaAlaAspGlyAlaGlnIysGlyGlyLysAsnIysSerSerGluAspAla 1634
Db 6536 CAGCTTCTGTTCAGCAGACAGCGGTGCCAGAAAGAGGTAAATTAAGTACAGAGATGCA 6595
QY 1635 LysSerValHisPheGluSerSerThrSerGluSerSerValLeuSerSerSerPro 1654
Db 6596 AAGCTGTCTATTTTGATTCACACACTCAGAGTCTCAGTGTATCAAGTATGATGCA 6655
QY 1655 GluSerThrLeuValIysProGluProArgGlyIleThrIleProGlyIleSerSerAsp 1674
Db 6656 GAGAGTACTTGTGTGAACAGAGCCGAATGGATTAACCTTCCTGTATCTTTCAGAT 6715
QY 1675 ValProGluSerAlaHisIysThrThrAlaSerGlyAlaIysSerAspThrGlyGlnPro 1694
Db 6716 GTCCAGAGAGTCCCAAACTATCTGCTCAGAGGCAAGTCAAGACTGTGGGAGCT 6775
QY 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714

Db 6776 ACCAAGTTGGACCTTTTCAGGTGACAACTACAGCAAAACAAAGTGGCTCTCTCTGTA 6835
 QY 1715 SerThrGluAspLysIleThrAspThrLysLysGluGlyProValAlaSerProPro 1734
 Db 6836 TCAAAAATGAGGACCAAGATACATGACACAAAGAAAGAGACAGTGGCATCTCTCTCT 6895
 QY 1735 PheMetAspLeuGluGlnAlaValLeuProAlaValIleProLysLysGluLysProGlu 1754
 Db 6896 TTATGATTTGGAAACAGCTGTTCTTCTGCTGTGATACCAAGAAAGAAAGAAACCTGAA 6955
 QY 1755 LeuSerLupProSerHisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSer 1774
 Db 6956 CTGTCACAGGCTTTCACATCTTAAATGGCCCTCTTCTACCCGAGGCGCTTTTATAGT 7015
 QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
 Db 7016 AGGATGTGATGATGTTGTTCCGATGATCCACACTGCCCTCAGCTGAGCTCAAAAGAC 7075
 QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSer 1814
 Db 7076 CTTCCTAGCCAGATCTAAGTCAAAAGCCTTAGTATTCATTACTCTCTCTTACATGAGT 7135
 QY 1815 SerAspAsnGluSerAspIleGluAspGluAspLeuLysLeuGluLeuAlaGlyLeuAla 1834
 Db 7136 AGCACAATAGTACATGATCGAAGATGAAACATTAAGTGAAGCTGCCAGCACTACGA 7195
 QY 1835 AspLysHisLeuLysGluIleGlnAspLeuGlnSerArgGlnLysHisGluIleGluSer 1854
 Db 7196 GATTAACATCTCAAAAGATTCAGAGACTGCAAGAGTCCGCAAGAGCATGAATTAATCT 7255
 QY 1855 LeuThrThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu 1874
 Db 7256 TTGTATACCAAACTGGGCAAGGTGCCCTCTGTATTATTTATTTCCCGAGCTGCCCTT 7315
 QY 1875 SerGlyArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
 Db 7316 TCAGGGAAGAGACACACACCCATTAAGCAAGAGCAAAATATGATCGAAGAGATTC 7375
 QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
 Db 7376 TTGGGGAATTAAGACCCCGACCTTTCAGGTACCTGCTGTGATGAGTGCAGCTTCAGTC 7435
 QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGluSerGlyGlnAsnGln 1934
 Db 7436 TTGCACCCCGACGAGACCCCTCCACCTCTGCAACATCCAGAGTCCGGGACAGATCAG 7495
 QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
 Db 7496 CTGTTCACAGCCCTTAAGCCATCTCCCTCCAGTACAAACCTCTATTCAGCCCTTACAGT 7555
 QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
 Db 7556 GATGTGTCCATTTCAAGTACCAAGCCTTCTCTCTCAGGTCAAGAAACCAAGACAGCAAAAC 7615
 QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSer 1994
 Db 7616 ACTGTGGGGCAACAGTAAACAGCAAGCCGCCCAAGCTCAGCTCCGCATGACGTC 7675
 QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLysLeuValAspAspThrAlaArg 2014
 Db 7676 AGCGAGGAAGGCAATTCACAGATGACTGCAAGTGTGAGCAATTTGGGCCCGAGAT 7735
 QY 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnLysGluGlyProGly 2034
 Db 7736 GCCATGATCTCTCTAGCGAGAGAGAGAGAAAGGCACTGAATTTAGAGAGGCGCTTCGA 7795
 QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLysCysIleSerMetThrSerAsnLeuGly 2054
 Db 7796 ATGGCAGGAAGATTTCTGACACTGGGCACTGTGACATCTCCATGACTCCAACTGGGT 7855
 QY 2055 GlySerAlaProLysSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMet 2074
 Db 7856 GGCTCTGCCCATCTCTGACAGCATCAAGTACTCTCTAGGTCACTTCAACCAAGTCTATG 7915

QY 2075 CysProGlnGlnLysGlyPheProAlaThrProPheGlyAlaGlnLysSerGlyThr 2094
 Db 7916 TGCCCCCAGACAGATGATGCTTTCAGTACCCCAATTTGGCGCTCATGAGAGGGACG 7975
 QY 2095 GlyGlyProAlaProLysProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGln 2114
 Db 7976 GTGGGCCACACACACAGCCACTGGCTGCACTTCCAACTGTGGGACCTGCTCTTCAG 8035
 QY 2115 AsnPheAsnLysSerAsnLeuGlnLysSerLysSerAsnProProGlySerAsnLeuArg 2134
 Db 8036 AATTTCACATGACGAAATTTGACAAATTCATGACAGACCCCGAGCTCCAACTGGG 8095
 QY 2135 ThrThr 2136
 Db 8096 ACCACT 8101

RESULT 5
 US-10-767-471-450
 ; Sequence 450, Application US/10767471
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001505
 ; CURRENT APPLICATION NUMBER: US/10/767,471
 ; CURRENT FILING DATE: 2004-01-30
 ; NUMBER OF SEQ. ID NOS: 50231
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ. ID NO: 450
 ; LENGTH: 8678
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(8678)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
 US-10-767-471-450

Alignment Scores:
 Pred. No.: 0 Length: 8678
 Score: 10604.00 Matches: 2125
 Percent Similarity: 89.21% Conservative: 0
 Best Local Similarity: 89.21% Mismatches: 11
 Query Match: 98.08% Indels: 246
 Gaps: 1

US-10-010-720-14 (1-2136) x US-10-767-471-450 (1-8678)

QY 1 MetSerGlyGlyAlaAlaGluLysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 956 ATGTCTGGCGCGCGCGAGAGAGACAGACAGCTCCGGTTCCGTCTCTCTCGCGC 1015
 QY 21 ProAlaProAlaProLysAsnGlySerSerSerAspSerValGlyLysLeuGly 40
 Db 1016 CCGGCTCTCTCCCGCAAGATGCTCCAGCTCCGATCTCCGTGGGAGGAACTGGGA 1075
 QY 41 AlaAlaAlaAlaAspAlaValThrGlyArgThrGluGlnLysArgArgArgHisThr 60
 Db 1076 GCCGCGCGCGCGAGCTGTGACCGGAGACCAAGAGTACAGCGCGCGCGCACACT 1135
 QY 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrThrGluHisArgPhe 80
 Db 1136 ATGACAAAGAGACACCGTGGGCGCGCGACCACTACCACTGAGACCGCTTCTTC 1195
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 Db 1196 CGCCGAGCGCTCATCTGACTCCAAATGCCACTGAGGCTTCCCGGCTTCTCTT 1255
 QY 101 SerLeuProGlnProSerLysProAlaAlaValProGlnSerAlaProProGluProHis 120
 Db 1256 TCCCTGCCCGACCCACGATCCCGCGGCTGTCCGAGAGTCTCCACCGGACCCGAC 1315

QY 121 ArgGluGluThrValThrAlaThrSerGlnValAlaGlnGlnProProAlaIa 140
 DB 1316 CGGAAAGAGACCGGACCGCCACCTTCCAGAGTACCCAGACCTCCAGCCCT 1375
 QY 141 AlaAlaProGlyGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 DB 1376 RCGCCCGCCGGGGAACAGAGCCGCGCGGGCCCTGCCCTCGACTGTCCCGAGGAGTACC 1435
 QY 161 SerIysAspArgProValSerGlnProSerLeuValGlySerIysGluGlnProProPro 180
 DB 1436 AGCAAAAGCCCGCCAGAGTCCAGCTTAGCTTGGGGGACCAAGAGACCCCGCCG 1495
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaIysGlnGluGluGluGluGluGluGluGlu 200
 DB 1496 GCGAAGATGGCAGCGCGCGCGCGCGCCCAAGAGCCACAGAGGAACGAGCCAGCAG 1555
 QY 201 GlnAspAspIleGluGluGluGluThrIysAlaValGlyMetSerAsnAspGlyArgPhe 220
 DB 1556 CAGAGTATATCGAAGAGCTGAGACCAAGCCGCGGGAATGTCTAAGATGCGCGCTTT 1615
 QY 221 LeuIysPheAspIleGluIleGlyValArgIysSerPheIysThrValIlyIysGlyLeuAsp 240
 DB 1616 CTCAAGTTTGCATCGAATCGGAGAGGCTCTTAAAGCGCTTCAAAAGGCTCGAC 1675
 QY 241 ThrGluThrThrValGluValAlaIleTrpCysGluLeuGlnAspArgIysLeuThrIysSer 260
 DB 1676 ACTAAACCAACGAGGAAGTGGCTGGTGTGATCGACGATCGAAATTAACAAAGCTCT 1735
 QY 261 GlnArgGlnArgPheIysGluGluAlaGluMetLeuIysGlyLeuGlnIleProAsnIle 280
 DB 1736 GAGAGGACGAGATTTAAAGAAAGAGCTGAATGTTAAAGCTTCACGATCCCAATATT 1795
 QY 281 ValArgPheIysAspSerTrpGluSerThrValIysGlyIysCysIleValLeuVal 300
 DB 1796 GTTAGATTTTATGATCTCGGGAATCCACAGTAAAGAAAGAGTGAATGTTTGGTG 1855
 QY 301 ThrGluLeuMetThrSerGlyThrLeuIysThrIysLeuIysArgPheIysValMetIys 320
 DB 1856 ACTGAACCTTATGACGCTCGGAACCTTAAACGATTCGAAAGGTTTAAAGTATGATAG 1915
 QY 321 IleIysValLeuAspSerTrpCysArgGlnIleLeuIysGlyLeuGlnPheLeuIleThr 340
 DB 1916 ATCAAAAGCTTCAAAAGCTGTGCGCTGAGATCTTAAAGGCTTCATCTTCATACT 1975
 QY 341 ArgThrProProIleIleHisArgAspLeuIysCysAspAsnIlePheIleThrGlyPro 360
 DB 1976 CGAACTCACTATTCATTCACCGGATCTTAAATGTAACAACATCTTATACCGGCTCT 2035
 QY 361 ThrGlySerValIysIleGlyAspLeuGlyLeuAlaThrLeuIysArgAlaSerPheAla 380
 DB 2036 ACTGGCTCAGTCAAGATTGGAGACCTCGCTGCGCAACCCGGAAGCGGGCTCTTTGGC 2095
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetIlyGluGluIysIysAsp 400
 DB 2096 AAGAGTGTAGTACCCAGAGCTTCAAGCTCCCTCAGATGTATGAGGAATATGAT 2155
 QY 401 GluSerValAspValIlyIalPheGlyMetCysMetLeuGluMetAlaThrSerGluThr 420
 DB 2156 GAATCGTTGAGTTTATGCTTTGGATGTGATGCTTGAGATGCTTACATCTGAATAT 2215
 QY 421 ProIysSerGluCysGlnAsnAlaIleGlnIleIlyIysArgValIlyThrSerGlyValIys 440
 DB 2216 CTTACTCGAGTCCCAAAATGCTGCGACATCTACCGTGGCGGACAGCGGGGTAG 2275
 QY 441 ProAlaSerPheAspIysValAlaIleProGluValIysGluIleIleGluGlyCysIle 460
 DB 2276 CCGACCGAGTTTGAACAAGTAGCAATTCCTGAAGTGAAGGAATATATGAAGATCAT 2335
 QY 461 ArgGlnAsnIysAspGluArgIysSerIleIysAspLeuLeuAsnIleAlaPhePheGln 480
 DB 2336 CGACAAACAAAGAGTGAAGATATTCATCAAAAGCTTTTGAACCAATGCTTTTCCAA 2395
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyIlyIysIleAlaIle 500

DB 2396 GAGAAACAGAGTACGGGTGAATTTAGACAGAAAGATGAGAAATATACCATTA 2455
 QY 501 LysLeuTrpLeuArgIleGluAspIleIysIysLeuIysGlyIysIysIysAspAsnGlu 520
 DB 2456 AAATTATGGCTACGATATTGAAGATATTAAAGAAATTAAGGAAAAATACAAAGATATGAA 2515
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet 540
 DB 2516 GCTATTGAGTTTCTTTTGTATTAGAGAGATGTCCAGAAAGATGTTGCAAGAAATG 2575
 QY 541 ValGluSerGlyIlyIalCysGluGlyAspHisIysThrMetAlaIysAlaIleIysAsp 560
 DB 2576 GTAGAGTCTGGATATGTCTGAGAGGTGATACAGACCATGCTTAAAGCTATCAAGAC 2635
 QY 561 ArgValSerLeuIleIysArgIysArgGluGlnArgGlnLeuValArgGluGluGln 580
 DB 2636 AGACTATCATTTAATTAAAGAGAAACGAGACGCGGAGTTGTATCGGAGAGACAGAA 2695
 QY 581 LysIysIysGlnGluGluSerSerLeuIysGlnGlnValGluGlnSerSerAlaSerGln 600
 DB 2696 AAAAAAAGCAGAAAGAGAGCATCTCAACAGAGGTAGAACATCCAGTGTCCAG 2755
 QY 601 ThrGlyIleIysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 DB 2756 ACAGGATCAAGCAGCTCCCTTCTGCTAGCACCGGCATCTACTGCTTACCACTTCA 2815
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnIleGlnIleu 640
 DB 2816 GCTTCAGTTTCTACACAAGTAAAGACTGAAGAACTGAGGACGATCAACATCAACACTA 2875
 QY 641 GlnTrpGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 DB 2876 CAGTACAGCAACCCAGATATATCTGTGTATCTGATGAGGAGCTTGAACAGGTGACAGGA 2935
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerIlyIysGln 680
 DB 2936 TCCTCTGCTTCAVAGATCTCGAGTGAAGACCAACAGCAAGTTTATATGTTCCCA 2995
 QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 DB 2996 CATGAACAGCACATCTTACAGGACAGCTCCAGGGCATVAACTTACTGTCCAGCA 3055
 QY 701 GlnSerGlnProHisGlyValIlyIysProProSerSerValAlaGlnGlyIysGlnGly 720
 DB 3056 CAGTCTAGCCCAAGTGGGTATATCCACCTCAAGTGTGACAGGGGACAGCCAGAGT 3115
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnIleProGln 740
 DB 3116 CAGCATCTCCAGTACTTAAACAGGGGTTTCACTTCCCAACCAACCAACATCTCAG 3175
 QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnIlySerLeuSer 760
 DB 3176 CAGCAGCAGGGAATACAGCAGACAGCCCTCTCAACAGACAGAGTATTCACCTTCA 3235
 QY 761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 DB 3236 CAGACATCAACCTCCAGTGAAGCCATCTACGACAGCAGCAGGATGATCCCTCAAGCTCA 3295
 QY 781 GlnValLeuProGlnValSerAlaGlyIysGln 791
 DB 3296 CAAGTCTTGCTCAAGTATCAGGTGGAAGAAACAGCTTCCAGTTCCAGCAGTACCACT 3355
 QY 791 ----- 791
 DB 3356 ATCCAAAGCGAAGCTCAGATCCAGTTGCGACACAAACCTCGTTGTTCCAGTCACTCT 3415
 QY 791 ----- 791
 DB 3416 GGTGCTCATTTCTCTCCAGTGGAGACAGCGCTCCCTATCTCCCTTGCTCCTCAGTACCT 3475
 QY 791 ----- 791

Db 3476 GTCTCTGATATCCCATATCACTCTCATGTGTACGCTCAGACAGGTTTCTCATCC 3535
QY 791 ----- 791
Db 3536 CTTCCTCATCATATGGCAGCTGGCATTAATCAGGCTCTGCTCAGTTGGCTTCACTGCT 3595
QY 791 ----- 791
Db 3596 ACAACAGCTGGGATCCGGGGGTATCAACTGTGTCTTAGTCAGCTTCCAAACCTTCTG 3655
QY 791 ----- 791
Db 3656 CAGCCTGTGACTCAGCTGCCAAGTCAGTTCAACCAAGCTCTTAACAACGACGATTGAG 3715
QY 791 ----- 791
Db 3716 TCCATGGGAATACCAAGCTAATCTTGACAAGCTGTGAGTTCCACTTCTCTGAGAT 3775
QY 791 ----- 791
Db 3776 GTTCTGTACAGGGCTTCCCACTCGACTGCCACACAGTACCCAGAGATTCAATATT 3835
QY 791 ----- 791
Db 3836 GCTCCCTCTTCCAACTGGCTTCTGTTTGCATTCATTCTACAGTCTTATCCCTCCCATG 3895
QY 791 ----- 791
Db 3896 CCGACAGAGTACTGGCTACACCTGGGTATTTCCACAGTGGTGCAGCCTTATGTGAA 3955
QY 791 ----- 791
Db 3956 TCAAACTTTTAACTTCTTATGAGTGTGTAGAGAGAGGTTCAAGTGTCCAGCCAGCA 4015
QY 792 ----- SerThGln 794
Db 4016 GGGAGTTTACACAGCCCCCACTACATCTCCACAGACAGTTTGGAGAGTACTCAG 4075
QY 795 G1YValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
Db 4076 GAGGTCTCTCAGGTGTCTCTGACAGCCAGTTGCCATACACAGMCCCAAGCTACCCAG 4135
QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
Db 4136 CCGACCACTTTGGCTCTCTGTAGACAGTCAATCATGTTGCTCAGGATAGT 4195
QY 835 AspG1YAsnGluAsnValProSerSerSerG1YArgHisGluGluThrSerArgProLysLeu 854
Db 4196 GATGGCATGAGAAAGTCCCATCTTCCAGTGAAGGATGAGAGAAAGAACTTACAAAACGG 4255
QY 855 HisTyrArgLysSerValArgSerArgSerArgHisGluLysThrSerArgProLysLeu 874
Db 4256 CATTAACGAAATCTGTAAAGAGTGGCTCTCGACATGAAAAAACTTACACCCCAAAATTA 4315
QY 875 ArgT1LeuAsnValSerAsnLysG1YAspArgValAlaGluCysGlnLeuGlnThrHis 894
Db 4316 AGAATTTTGAATGTTTCAAAATAAAGGACCGAGTGTAGATGATCAATTAAGAGACTCAT 4375
QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspG1YAspAsnProGluGluLeuAla 914
Db 4376 AATAGAAATGTGTATCATTAATTTGACCTAGATGTGACAAACCCCGAGGAGATGCA 4435
QY 915 ThrT1MetValAsnAsnAspPheT1LeuAlaT1LeuLysGluSerPheValAspGln 934
Db 4436 ACAATATATGTGAACATGACTTTATCTAGCAATAGAGAGAGAGTGTGTTGTGATCA 4495
QY 935 ValArgGluT1LeuT1LeuLysValAspG1YMetLeuSerGluAspValSerValGluPro 954
Db 4496 GTGCGAAGAAATTATGAAAAAGCTGAAATGCTCAGTAGAGATGCTCACTGGAACCA 4555
QY 955 G1UG1YAspG1YAsnGluLysGluSerLeuGlnG1YLysAspAspTyrG1YPheSerG1YSer 974
Db 4556 GAGGTGATCAAGGATTTGGAGAGTCTACAGAGAAAGATGACTATGGCTTTTACAGTTCT 4615

QY 975 G1YLeuGluGluGluPheLysGlnProT1LeuProT1LeuSerSerMetProGlnGlnT1Leu 994
Db 4616 CAGAAATTGAGAGAGAGTTCAAAACCAATTCCTGCTTCCATGACCAAGCAAAATA 4675
QY 995 G1Y1LeuProThrSerSerLeuThrGlnValAlaHisSerAlaG1YArgArgPheT1LeuAl 1014
Db 4676 GGATTCCTTACAGATTCTTAACTCAAGTTGTTCACTCTGGGGAGGCGGTTATATGTT 4735
QY 1015 SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluT1LeuThrAsp 1034
Db 4736 AGTCTGTGCCAGAAAGCCGATTACAGAAATCAAAAGTTTTCCTCCAGTGAATTAACAGAT 4795
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProG1YMetAsnLeuSerHisSerAlaSerSer 1054
Db 4796 ACAATTCTCTCTTACAGCTCAGAGAGCTCGAATGAACTTGTCTCATCTGCTCATCTCC 4855
QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGlnMetThrGluG1YProAsn 1074
Db 4856 CTTAGTCTACACAGGCTTTTCTGAACCTTAAAGAGTGCACCAATTCAGTAGTACCTCTTCTTA 4915
QY 1075 ThrAlaProAsnPheSerHisThrG1YProThrPheProValValProPhePheLeu 1094
Db 4916 ACAGCACTTCAAACTTTAGTCAATACAGAGCAACAACTTCCAGTAGTACCTCTTCTTA 4975
QY 1095 SerSerT1LeuAlaG1YValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
Db 4976 AGTAGCATTCCTGAGTCCCACTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 5035
QY 1115 SerProAsnAspLysSerThrSerSerVal1LeuSerGluValThrValProThrGlu 1134
Db 5036 AGCCTCTCTTAAATGACATTTCCATCATCAGTATTCAGTCTAGGTTACAGTCCCATGAA 5095
QY 1135 G1UG1Y1LeuAlaG1YValAlaThrSerThrG1YValValThrSerG1YLysProT1Leu 1154
Db 5096 GAGGGATGTCTGAGTGTGCACACAGCAGCAGGTTGTGTAACTTCAAGTGTCTCCCATTA 5155
QY 1155 ProProValSerGluSerProValLeuSerSerValValSerSerT1LeuT1LeuProAla 1174
Db 5156 CCACTGTCTGTAATACCAACAGTACTTCCAGGAGTATTTCAAGTATACAAATACCTGCA 5215
QY 1175 ValValSerT1LeuSerThrThrSerProSerLeuGlnValProThrSerThrSerGluT1Leu 1194
Db 5216 GTTGTCTCATATCTATCATCATCCCGTCACTTCAAGTCCCAATCACAATCATGTGATGC 5275
QY 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 5276 GTTGTCTTAAAGTACAGACATGATCTTCAAGTAAAGTTTCAAGCAACTTCAAGCCTCTGCA 5335
QY 1215 G1YG1YSerThrAlaThrProGluProLysProProAlaValAlaSerGlnGlnAlaAla 1234
Db 5336 GGGGGAGATGCTCTACCCAGGCTCTTACGCTCCAGCTGTAGTATCTCGAGGAGCA 5395
QY 1235 GlySerThrThrValG1YAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 5396 GGCAGACACTAGTGGAGGCACTTAAATCATCAGTTTCAACCACTTCAATCCCAAGC 5455
QY 1255 ThrAlaSerGluLeuSerT1LeuGlnLeuSerSerSerThrSerThrProThrLeuAlaGlu 1274
Db 5456 ACAGCTTCAAGCTGTGSCATTCAGTTCAGGAGAGTCACTTCACTCTTCAAGTCTGAA 5515
QY 1275 ThrValValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrG1YLeu 1294
Db 5516 ACCGTGTATGTAGCCACACTCATCAGTAAATACATCTCATGCACTCAATGATG 5575
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProG1YAlaG1YValSerSerTyr 1314
Db 5576 GCTTCTCTCTCTCTGACACATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5635
QY 1315 T1LeuSerGlnProG1YG1YLeuHisProLeuValT1LeuProSerValT1LeuAlaSerThrPro 1334
Db 5636 ATTTCAGCTGTGTGGCTGATCTCTTGTGTCATTCATCAAGTATGATGATCTTACTCTCT 5695

QY 1335 IleLeuProGlnAlaIleGlyProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 5696 ATTCTTCCCAAGCAGAGGAGCTACTTCTTACACCTTTATTACCCCAAGTACTAGTATC 5755
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 5756 CCACCTTGATGAGCTGTTGGCCATGTGCTGCTGTACAGAGACACTAATTCATAGT 5815
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValAspSer 1394
Db 5816 CAGCTTCACAGCTGTTGCTTCCACACAGCCCATCTCATGTCTCTGAGAGATATCT 5875
QY 1395 AspThrGlnProValAspProGlyIleAspAspIleThrLeuGlnGlnValLeuArg 1414
Db 5876 GATACACACCCCAAGCTCTGGAATTGATGACATTAAGACTCTGAGAAGAAAGCTGCG 5935
QY 1415 SerLeuPheSerGlnHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 5936 TCTCTGTTCAAGTACACACTCATCTGAGCTCAGATGCTCTGTCTCATCTGAGAACCC 5995
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 5996 TCACTAGTCAATAGAGACACTGTACACAGGCACTCCAACTACTGCTGTGACCCAAAGC 6055
QY 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 6056 AAACCTCTGACTTCTACACAAAGTACTGCTTACCAACCAATTTTACCACTAGAACAC 6115
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 6116 GTTGCTTTGACAGTTACACAGAGTGTACACCTGGGCAAGTTTCTACCCAGTACAGCAT 6175
QY 1495 ThrThrSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysValPro 1514
Db 6176 ACTACATCAGAGAGTGAACCTGGAACCTGCTCCCAAGCCCACTTCACTTAAAGCTCCGC 6235
QY 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 6236 GTGTGTGCAAGTGGGTACTGAACTTCCAGAGTACTTACCCAGCGAGAGCTGCCACT 6295
QY 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 6296 TTTCAGAGACTTCTCTAACCCAGTCCAGCAACCTCTAGAGATCTTGATGCTCAATTCG 6355
QY 1555 ArgArgThrLeuSerProGln**IleThrValThrSerAlaValGlyProValSerMet 1574
Db 6356 AGAAGAACACTTACTCAGAGATKATACAGTACAGTCTGTGGTGTGCTGTGTCATG 6415
QY 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 6416 GCGGCTCYAACAGCAATCACAGAACGAGAACACAGCTCAGAAAGGTGTTTCCCAAGTC 6475
QY 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 6476 AAAGAAAGCCCTGCTTACGAACTAGTTCAGAGACTGTGTTTAAAGTAGCGGACGATTT 6535
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnLysAla 1634
Db 6536 CAGGTTTCTGTGAGAGAGAGGTCAGAAAGAGGTAAATAAGTCAAGAAAGATGCA 6595
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerValLeuSerSerSerSerPro 1654
Db 6596 AAGTCTGTTTCTTATTTGATCCAGCACTCAGAGTCTCAAGTGCATACAGTAGTACGCCA 6655
QY 1655 GlnSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 6656 GAGAGTACTTGTGGAACCCAGAGCCCAATGCAATACATCCCTGATATCTTTCAGAT 6715
QY 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlnPro 1694
Db 6716 GTGCAAGAGGTGCCCAAAAATCTAGCTCAGAGGCAAAAGTACAGACTGTGGCAAGCT 6775
QY 1695 ThrIleValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 6776 ACCAAGTTGAGCTTTTTCAGCTGACACACTACAGCAACAAAGTGGTCTTCTCTGTA 6835
QY 1715 SerLysThrGlnLysLysIleThrAspThrLysLysGlyGlyProValAlaSerProPro 1734
Db 6836 TCAAAAATCTAGGACAAGATCACTGACACAAAGAAAGAGAACAGTGGCATCTCTCCT 6895
QY 1735 PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlyLysProGln 1754
Db 6896 TTTATGATTTGAAACAGCTGTTCTTCTGTGTGATACCAAGAAAGAAAGAAAGCTGAA 6955
QY 1755 LeuSerGlnProSerHisLeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSer 1774
Db 6956 CTGTCAGAGCTTCAATCAATCAATGAGCGCTCTTCTGACCCGAGAGCGCTTTTATAGT 7015
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
Db 7016 AGGATGTGATGATGATGTTCCGGTAGTCCACACTGCCCATCAGCTGAGCTCAAAAGAGC 7075
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSer 1814
Db 7076 CTTCCTAGCCAGAAATCTAAGTCAAAAGCTTAGTATTATTTAATCTCTTACTATAGT 7135
QY 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuGlnLeuArgArgLeuArg 1834
Db 7136 AGCGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7195
QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer 1854
Db 7196 GATTAACATCTCAAGAGATTCAGAGACTTCAGAGTGCAGAGAGCATGAAATTTGATCT 7255
QY 1855 LeuThrThrLysLeuGlnGlyLysValProProAlaValIleIleProProAlaAlaProLeu 1874
Db 7256 TTGTATATCCAAACTGGGCAAGTGGCCCCCTGCTGTATTTATTTCCCCAGCTGCTCCCTT 7315
QY 1875 SerGlyArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSer 1894
Db 7316 TCAAGGAGAAAGACAGACACCACTAAAGCAAGAGCAGCAATCTAGTCCAAAGAGATTC 7375
QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
Db 7376 TTGGGGAATTAAGACCCCGCTTCAAGTAACTGTGTGTGTGTGTGTGTGTGTGTGTGT 7435
QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
Db 7436 TTGCACCCCAGACAGACCTTCCACCTCTGCGCAACATCCAGATCGGGCAGAAATCAG 7495
QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuLysSerAlaPheThrSer 1954
Db 7496 CTGTTACAGCCCTTAAGCCATCTCCCTCAGTACCAACTCTATTCAGCTTACACAGT 7555
QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerThrAsn 1974
Db 7556 GATGTGTCATTTTCAGTACAGACCTTCTGCTTCAGAGTCAAGGAAACAGACAGCAAAAC 7615
QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnIleProProAlaMetThrSer 1994
Db 7616 ACTGTTGGGCAACAGTGAACAGCAAGCCGCCCAACTCAGCTCTGCAATACACTCC 7675
QY 2015 AlaMetAsnLeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrgLysProGly 2034
Db 7736 GCCATGATCTTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7795
QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
Db 7796 ATGGCAAGAGAGTCTTGTGACCTGGGCACTGTGCAATCTCCATGACTCGAAGCTGGGT 7855
QY 2055 GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlnHisPheThrLysSerMet 2074

QY 481 GlnGluThrGlyValArgValGluLeuAlaGluIleuAspAspGlyGluIleuValLeuIle 500
 Db 2396 GAGGAAACGAGAGTACGGGTGAATTAAGAGAAAGATGATGGAGAAATTAAGCATTA 2455
 QY 501 LysLeuThrPheArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
 Db 2456 AATATATGGCTCGATATTAAGATATTAAGAAATTAAGGAAATTAAGATATATGA 2515
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 Db 2516 GCTATTGATCTTCTTCTTATTAAGAGAGATGTCCAGAAATGTTCCACAGAAATG 2575
 QY 541 ValGluSerGlyTyrValAlcysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 Db 2576 GTAGAGCTGGGATATCTGTGAAGGTGATCACAAACCATGGCTTAAGCTATCAAGAGC 2635
 QY 561 ArgValSerLeuIleLysArgLysArgGluArgGlnArgGlnLeuValArgGluGluGln 580
 Db 2636 AGAGTATCATTTAATTAAGAGAAACGAGAGCAGCGCAGTTGGTACGGAGAGCAAGAA 2695
 QY 581 LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGluGluSerSerAlaSerGln 600
 Db 2696 AAAAAAAGCAGAGAGAGAGCAGCTCAACAGCAGTAGAACATCCAGTGTCTCCAG 2755
 QY 601 ThrGlyLysGlnLeuProSerAlaSerThrGlyLysProThrAlaSerThrThrSer 620
 Db 2756 AGGGAATCAAGAGATCTCTCTGTAGACCGGATACCTACTGCTTCTACCATTTCA 2815
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
 Db 2816 GCTTCAAGTTCTACACAGTAGAACCTGAAGAACCTGAGGAGCATCAACATCAACAATA 2875
 QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyTyrValAspSerGlyGlnGly 660
 Db 2876 CAGTACAGAGAACCCAGTATATCTGTATCTGATGGACGGTTCACAGTGTTCAGGAA 2935
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 Db 2936 TCTCTGTCTTTCAYAGATCTCAAGTAGAGCCACAGCAGCTTTCATATGTTCCCA 2995
 QY 681 HisGluGlnAlaHisSerThrGlyThrValProGluHisIleProSerThrValGlnAla 700
 Db 2996 CAGGAACAGGCAATCTCAAGGACAGCAGTCCAGGGCATATVACTTACTGTCCAAGCA 3055
 QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 Db 3056 CAGTCTCAGCCCATGGGGTATATCCACCTCAAGTGTGACACAGGGGAGAGCCAGGGT 3115
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
 Db 3116 CAGCCATCTCAAGTAGCTTAAACGGGGTTCATCTTCCACCCATACAAACATCTCCAG 3175
 QY 741 GlnGlnGlnGlnIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
 Db 3176 CAGAGAGAGGAAATACAGCAGACAGCCCCCTCTCAACAGACAGGAGTATTCATCTTCA 3235
 QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 Db 3236 CAGACATCAACCTCCAGTGAAGGCCACTACAGCAGCAGTAGAGTARCTCTCAAGCTCCA 3295
 QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
 Db 3296 CAAAGTCTGCTCAAGTATCAGCTGGAAACAGCTTCCAGTTTCCAGCCAGTACCAACT 3355
 QY 791 ----- 791
 Db 3356 ATCCAAAGCGAAGCTCAGATCCAGTTGCGACACAACTCGGTTTTCAGATCCACTCT 3415
 QY 791 ----- 791
 Db 3416 GGTGCTCATTTCTTCCAGTGGAGACAGCCGCTCCTTACTCCTTGTCTCCTCAGTACCT 3475
 QY 791 ----- 791

Db 3476 GTCCTCARTTCCCATATCAACTCTCATATGTGTACAGGCTCAGACAGGTTTCTCATCC 3535
 QY 791 ----- 791
 Db 3536 CTTCCCATCAACATGAGCAGCTGGCATTTACTACGCTCTGCTCAGCTTGGCTTCATCTGCT 3595
 QY 791 ----- 791
 Db 3596 ACAACAGCTGCATCCCGGGGTATCACTGTGGTTCTTACTAGTCTTCCAACTCTTCTG 3655
 QY 791 ----- 791
 Db 3656 CAGCCTGTGACTCAGTGGCAAGTCAGTTCAACCAACAGCTCTCTACACAGCAGATTGAC 3715
 QY 791 ----- 791
 Db 3716 TCCATGGGAATACAGCTAATCCTTGGACAGCTGTGAGTTCCACTTCTCTGTGAGAT 3775
 QY 791 ----- 791
 Db 3776 GTTCTGTACAGAGGCTTCCCACTGACTGCCACAGTACCCAGATCCAGATTCATATTT 3835
 QY 791 ----- 791
 Db 3836 GCTCCCTTCCACAGTGGCTTCTGTGTGATCATTCTTACAGTCCATCCCTCCCATG 3895
 QY 791 ----- 791
 Db 3896 CCGACAGAAAGTACTGGCTACACCTGGGATCTTCCACAGTGTGACAGCTTATGTGAA 3955
 QY 791 ----- 791
 Db 3956 TCAAATCTTTAGTTTCTTATGGTGTGTAGAGAGCAGGTTCAAGTGTCCAGCCAGGA 4015
 QY 792 ----- SerThrGln 794
 Db 4016 GGGAGTTTAGCAACAGCCCCCACTACATCTCCACAGAACAGATTGAGAGATCTACAG 4075
 QY 795 GlyValSerGlnValAlaProAlaGluProValAlaValAlaGlnProGlnAlaThrGln 814
 Db 4076 GGAGTCTCAGGTTCTCTCTGAGAGCCAGTGGCAGTGCAGACAGCCCAAGCTTACCCAG 4135
 QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 Db 4136 CCGACCACTTGGCTTCTCTGTGAGACAGTGCATTCAGATGTCTCTCAGGTATAGAT 4195
 QY 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGluGlyArgThrThrLysArg 854
 Db 4196 GATGCAATGAGAACCTCCATCTTCAGTGAAGGACATGAGAAAGAACTACAAAACGG 4255
 QY 855 HisTyrArgLysSerValArgSerArgSerArgHisGluLysThrSerArgProLysLeu 874
 Db 4256 CATTACCGAAATCTGTGAAGAGTGGCTCTCAGCATGAAATTAATTCAGCCCAAAATTA 4315
 QY 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValAlaGluCysGlnLeuGlnThrHis 894
 Db 4316 AGAATTTTGAATCTTCAATTAAGAGACCGAGTGTAGAAATGTCAATTAGAGACTCAT 4375
 QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGluGluIleAla 914
 Db 4376 AATAGGAAATGTGTTCAATCAATTTGACCTGATGTGTACAAACCCGAGGAGATAGCA 4435
 QY 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGluArgLysSerPheValAspGln 934
 Db 4436 ACAATTTATGTGAACAATGACTTATTTCTGCAATAGAGAGAGATGCTTTGTGAGATCA 4495
 QY 935 ValArgGluIleIleGluLysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
 Db 4496 GTGCGAAGAAATTTATGAAAAAGCGATGAATGTCTGATGATGTCACTGTGGAACCA 4555
 QY 955 GlnGlyAspGlnGlyLeuGluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974

Db 4556 GAGGGTATCATGAGATTGAGAGCTCTACAGAAAGATGACTATGGCTTTTCAGGTTCT 4615
Qy 975 GlnIleuGlnGlnGlnIleuPheGlnProIleProIleSerSerMetProGlnGlnIle 994
Db 4616 CAGAAATTTGAGAGAGATTCAAAACAACTTCTGCTTCTTCATGCGCACAGCAATA 4675
Qy 995 GlnIleProThrSerSerLeuThrGlnValValHisSerIleGlnValGlnPheIleVal 1014
Db 4676 GGCATTCTCTACAGATTCTTTAACTCAAGTTGTTCTCTGCGGAGAGCGGTTATAGTG 4735
Qy 1015 SerProValProGlnSerArgLeuArgGlnSerIleValPheProSerGlnIleThrAsp 1034
Db 4736 AGTCTGTGCGAGAAAGCCGATTCAGAAATCAAAAGTTTCCCAAGTGAATTAACGAT 4795
Qy 1035 ThrValAlaIleSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db 4796 ACAGTTCTGCTCTTACAGCTCAGAGCCCTGGATGAATCTTGTCTCATCTGCTCATCTCC 4855
Qy 1055 LeuSerLeuGlnGlnAlaPheSerGlnLeuArgArgIleGlnMetThrGlnGlyProAsn 1074
Db 4856 CTTAGCTACAAACAGGCGCTTTTCTGAACCTTAGAGCTCCCAATGACAGAAAGACCAAN 4915
Qy 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValValProProPheLeu 1094
Db 4916 ACAGACCTCTCAAACTTAGCTATACAGAGCAACCAATTTCCAGTAGAACCTCTTCTTA 4975
Qy 1095 SerSerIleIleGlnValProThrThrAlaAlaIleThrAlaProValProAlaThrSer 1114
Db 4976 AGTAGCATTTGTGAGTCCCAACACAGAGCAAGCCAGAGCCAGTCCCTGCAACAGC 5035
Qy 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGlnValThrValProThrGln 1134
Db 5036 AGCCCTCTCATATGACATTTCCACATCATGATTCAGTGAAGTTACAGTGCCACATGAA 5095
Qy 1135 GlnGlnIleIleGlnValAlaIleThrSerThrGlyValValValThrSerGlnGlyLeuProIle 1154
Db 5096 GAGGGAGATCTGAGTTGCGACAGCAAGGTGTGAACCTTCAGGTGTCTCCCATTA 5155
Qy 1155 ProProValSerGlnSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 5156 CCAGCTGTGTGATACACAGATCTTCCAGGCTATTTCCAGTATCACAAATACCTGCA 5215
Qy 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGlnIle 1194
Db 5216 GTTGTTCAATATCTACTACATCCCGCTCACTTCAGATCCCAATCATCATCATGAGATC 5275
Qy 1195 ValValSerSerThrAlaLeuArgProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 5276 GTTGTTCTAGTACAGCATGTATCTTCAAGTAAACAGTTTCAGCAACTTCAGCCTTGCA 5335
Qy 1215 GlnGlySerThrAlaThrProGlyProGlyProProAlaValValSerGlnGlnAlaIle 1234
Db 5336 GGGGGGAGTACTGTACCCAGAGCTTAAGCTTCCAGCTGTAGTATCTGACAGAGCA 5395
Qy 1235 GlySerThrThrValGlnAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 5396 GGGAGCATCTAGTGGAGGACATTAACATGATTTCTACACACCACTTATCCCAAGC 5455
Qy 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrIleAlaGln 1274
Db 5456 ACAGCTTCACAGCTGTCATTCAGTTAGAGCAAGTACTTCACTCTTACCTTAAAGTGAA 5515
Qy 1275 ThrValValValSerAlaHisSerLeuAspIleThrSerHisSerSerThrThrGlnLeu 1294
Db 5516 ACCGTGTACTTACCGCACACTCATAGATTAAGACATCTCATAGCAGTAAACCTGATG 5575
Qy 1295 AlaPheSerLeuSerAlaProSerSerSerSerProGlyAlaGlyValSerSerTyr 1314
Db 5576 GCTTTCTCCCTCTGACACATCTTCTCTCTCTCTGAGAGAGAGTCAAGTAT 5635
Qy 1315 IleSerGlnProGlnGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
Db 5636 ATTTCAGGCTGGTGGGCTGACATCTTTGATCATTCATCAGTAAAGCTTACTCTCT 5695

Qy 1335 IleLeuProGlnAlaIleGlnIleProThrSerThrProLeuLeuProGlnValProSerIle 1354
Db 5696 ATTCTTCCCAAGAGAGAGACCTACTCTTACACCTTATTAACCCCAAGTACCTAGATC 5755
Qy 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 5756 CCACCTTGTGTACAGCTGTGGCAATGTCTGCTGTACAGCAGACATTAATCATAGT 5815
Qy 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValAspSer 1394
Db 5816 CAGCTTCACAGCTTTGTTCTCCCAACCAACCCCAATCACTATGTCTGGAAGTAACTCT 5875
Qy 1395 AspThrGlnProValAspProGlyIleAspAspIleThrThrLeuGlnGlnValLeuArg 1414
Db 5876 GATACACACCCAAAGCTCTGGAATTTAGACATTAAGACTCTAGAAAGAAAGCTGGG 5935
Qy 1415 SerLeuPheSerGlnHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 5936 TCTCTGTTCAGTGAACACAGCTCATCTGAGCTCAGCATGCTGTCTCACTGAGAGCC 5995
Qy 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 5996 TCAGTGTCAATGAAGAGACAGCTGTACACAGGATCCCACTACTGCTGTGACCAAGC 6055
Qy 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 6056 AAACCTCTGACTTCTACCAAAAGTACTGCTTACCAACCAACCAATTTACCACTAGAAC 6115
Qy 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 6116 GTTGCTTTGCACTTACACAGAGTGTACACCTGGGCAAGTTCTTACCCAGTACAGACT 6175
Qy 1495 ThrThrSerGlyValLysProGlyYThrAlaProSerIleYsProProLeuThrIleValPro 1514
Db 6176 ACTACATCAGAGATGAAACCTGGAACCTGCTCCCTCCAAAGCACCTCTAACCTAAAGCTCCG 6235
Qy 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 6236 GTGCTGCAAGTGGTACTGAACTTCCAGAGTATCTTACCAGCAGCAGCTGCCACT 6295
Qy 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 6296 TTTCCAGAGACCTTCTTAACCCAGTCCCAACACTCTTAAGAGATCTTGATGCTCAATTG 6355
Qy 1555 ArgArgThrLeuSerProGlnU***IleThrValThrSerAlaValGlyProValSerMet 1574
Db 6356 AGAAGAACACTTAGTCCAGAGATATACAGTACTTGCGGTGGTCTGTGTCATG 6415
Qy 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnIleGlyValSerGlnVal 1594
Db 6416 GCGGCTCYAACACACATCAKRAAGCAAGAAACACAGCTCAGAAAGGTGTTCTCAAGTC 6475
Qy 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheIleYsMetGlyArgPhe 1614
Db 6476 AAAGAAAGCCCTGCTCTAGCAACTAGTTCAGAGAGTGTGTTTTAAAGTGGAGCATTT 6535
Qy 1615 GlnValSerValAlaIleAspGlyAlaGlnIleGlnGlyLysAsnIleYsSerGlnIleAspAla 1634
Db 6536 CAGGTTTGTGTACAGAGCGGTGCCAAGAAAGAGGTAAATAGTCAAGAAAGCA 6595
Qy 1635 LysSerValHisPheGlnSerSerThrSerGlnSerSerValLeuSerSerSerPro 1654
Db 6596 AAGTGTTCATTTTGAATCAACACCTCAGATCTCAGTGTCTATCAATAGTATAGTCA 6655
Qy 1655 GluSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 6656 GAGAGTACCTTGGTGAACAGAGCCGAATGGCATAAACATCCCTGGTATCTCTTCAAT 6715
Qy 1675 ValProGlnSerAlaHisIleThrThrThrAlaSerGlnAlaLysSerAspThrGlnPro 1694
Db 6716 GTGCCAGAGAGTCCCAAAACCTACTGCTCTCAGAGCAAGTCAAGCACTGGGAGGCT 6775

QY	1695	ThrlYsValGlYAcPpHeGInValThThrThrAlaSnYsValGlYAcPpHeSerVal	1714
Dp	6776	ACCAAGGTGGACGTTTCAGGTGACACTACACAAACAAAGGGCGTTCTCTGTGA	6835
QY	1715	SerYsThrGlnAspLysIleThrAspThrYsIysGlnGlyProValAlaSerProPro	1734
Dp	6836	TCAAAACTGAGGACCAAGATCACTGACACAAAGAAAGAGGACCGATCTCTCTCT	6895
QY	1735	PheMetAspLeuGlnGlnAlaValLeuProAlaValIleProYsIysGlnLysProGln	1754
Dp	6896	TTTATGTGATTGTGGAAACAAGCTGTTCTCTCTCTGTATACCAAAAGAAAGAAACCTCGAA	6955
QY	1755	LeuSerGlnProSerHisIleuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSer	1774
Dp	6956	CTGTGACGCGCTTCACATCTAAATAGGGCGGCTCTCTGACCGGGAGCGCTTTTATAGT	7015
QY	1775	ArgAspValAlaAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer	1794
Dp	7016	AGGAGATGGATGATGATGTTCCGGTATGTCACACTCGCCCATGACTGACTCAAAAGC	7075
QY	1795	LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyrMetSer	1814
Dp	7076	CTTCTCAGCCAGAACTCTAAGTCAAAAGCCTTGATTCATTATTAACCTCTTACATGAGT	7135
QY	1815	SerAspAsnGlnSerAspIleGlnAspGlnAspLeuYsLeuGlnLeuAlaArgLeuArg	1834
Dp	7136	AGCGACAAATGATGATGATATGAAAGATGAAACTTAAATGATGAGCTGGCAGCATGCA	7195
QY	1835	AspLysHisIleuLysGlnIleGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSer	1854
Dp	7196	GATTAACACTCTCAAAAGAGATTCAGGACCTGACGAGCTGCCCAAGACATGAATGAATCT	7255
QY	1855	LeuTyrThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu	1874
Dp	7256	TTGTATACCAAACTGGGCAAGTGGCCCCCTGCTGTATTATTATCCCCAGCTGCCCTT	7315
QY	1875	SerGlnAlaArgAlaArgArgProThrLysSerYsGlySerLysSerSerHisSerSerSer	1894
Dp	7316	TCAGGGAACAAACACACGACCCACTTAAAGCAAAAGCACAAATCTAATGCGAAGAGATGCC	7375
QY	1895	LeuGlnAlaLeuLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal	1914
Dp	7376	TTGGGGAAATAAAGCCCCCAGCTTTCAGGTAACTCTGTGTGATGAGTGACGCTTCAAGTC	7435
QY	1915	LeuHisProGlnGlnThrLeuHisIleProProGlyAsnIleProGlnSerGlyGlnAsnGln	1934
Dp	7436	TTGCACCCCCACGACAGCCTCCACCTCTGCGCAACATCCGACGAGTCGGGCAAAATCAG	7495
QY	1935	LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuTyrSerAlaPheThrSer	1954
Dp	7496	CTGTATACGCCCCCTTAAACCACTCTCCCTCCAGTGCACAACTCTTATTCAGCTTCAACAGT	7555
QY	1955	AspGlnAlaIleSerValProSerLeuSerAlaProGlyGlnGlyYlnThrSerSerThrAsn	1974
Dp	7556	GATGTGTCATTTCAGTACCAAGCCTTTCCTCTCAGGTCAAGGAACGACGACCAAAAC	7615
QY	1975	ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSer	1994
Dp	7616	ACTGTGGGGGCAACAGTGAACAGCCAAAGCGGCCCAAGCTCAGCCTCGGCATGACGTCC	7675
QY	1995	SerArgLysGlyThrPheThrAspAspLeuHisIysLysLeuValAspAspTTPAlaArgAsp	2014
Dp	7676	AGCGAGGAAGGCGACATTCACAGATGACTTGCAACAAGTTGTATGACAATTTGGCGCCGAGAT	7735
QY	2015	AlaMetAsnLeuSerGlyArgArgGlySerYsGlyHisMetLeuTyrGlnGlyProGly	2034
Dp	7736	GCCATGAATCTCTCAGGCAAGGAGGAAAGCAAGGCGACATGAATTANGAGGGCGCTTGGGA	7795
QY	2035	MetAlaArgLysPheSerAlaProGlyGlnLeuYsIleSerMetThrSerAsnLeuGly	2054
Dp	7796	ATGGCAAGGAAGTCTCTGCAACCTGGGCAACTGTGCATCTCCATGACCTCGAACCTGGGT	7855
QY	2055	GlySerAlaProIleSerAlaAlaSerAlaThrSerLeuGlnYHisPheThrLysSerMet	2074

```

Db      7856 GAGCTGCCCCCATCTCTGACATCATGCTACTCTCTAGTACTTCAACCAATCATG 7915
QY      2075 CysProProGInGInTyrgLyPheProAlaIarhProPheGlyAlaGInTyPserGlyThr 2094
Db      7916 TGCCCCCACAAGACAGTATGAGCTTTCCAGCTACCCCATTTGGGGCTCAATGAGTGGGACG 7975
QY      2095 GlyGlyProAlaProGInProLeuGlyGInPheGInProValGlyThrAlaSerLeuGIn 2114
Db      7976 GGTGGCCCAAGACCAACAGCATCTTGCCAGTTCCAACTGTGGGAACTGGCTCTCTGACG 8035
QY      2115 AsnPheAsnIleSerAsnLeuGInTyPserIleSerAsnProProGInTyPserAsnLeuArg 2134
Db      8036 AATTTCACATCATGCAATTTCGAGAAATCATCATGCAACCCCAAGGCTTCAACTTGGGG 8095
QY      2135 ThrThr 2136
Db      8096 ACCACT 8101

RESULT 7
US-10-767-471-455
; Sequence 455, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 455
; LENGTH: 10464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(10464)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-2
US-10-767-471-455

Alignment Scores:
Pred. No.: 0 Length: 10464
Score: 10604.00 Matches: 2125
Percent Similarity: 89.21% Conservative: 0
Best Local Similarity: 89.21% Mismatches: 11
Query Match: 98.08% Indels: 246
DB: 7 Gaps: 1

US-10-010-720-14 (1-2136) x US-10-767-471-455 (1-10464)
QY      1 MetSerGlyGlyAlaAlaGlnTyPserIleSerThrProGlySerLeuPheLeuSerPro 20
Db      956 ATGTCTGGCGGGCGCCGACAGAGAAGCAGAGACACACTCCCGGTTCCCTCTCTGCGCG 1015
QY      21 ProAlaProAlaProLyAsnGlySerSerSerAspSerSerValGlyGlnTyPserLeuGly 40
Db      1016 CCGGCTCTCTGCCCCCAAGATGCTCCAGCTCCGATCTCTCCGCTGGGGGAGAAACTGGGA 1075
QY      41 AlaAlaAlaAlaAspAlaValThrGlyArgThrGlnGlnTyArgArgArgArgHisThr 60
Db      1076 GCCGGGGCGCCGAGAGCTGTGACCGGCAGAGACCGAGAGTACAGGGCGCGCCGACACT 1135
QY      61 MetAspLyAspSerArgGlyAlaAlaIaIaThrThrThrThrGlnHisArgPhePhe 80
Db      1136 ATGGACAAAGAGAACCGCTGGGGGGCGCGGACCACTACACACACTAGACACCGCTTCTTC 1195
QY      81 ArgArgSerValIleCysAspSerAsnAlaIaIaLeuGlnLeuProGlyLeuProLeu 100
Db      1196 CGCCGAGGAGTATCTGACATCCCAATGCCACTGCACCTGGAGCTTCCGGGCTTCTCTT 1255
QY      101 SerLeuProGInProSerIleProAlaAlaValProGInSerAlaProProGInProHis 120

```

Dh 1256 TCCTGCCCCAGCCAGCATCCCGCGGCTGTCCGAGAGTGTCTCCAGCCGAGCCGCCAC 1315
Qy ArgLugluThrValThrAlaThrSerGlnValAlaGlnGlnProProAla 140
Dh 1316 CGGAGAGAGACCGGAGACCGCCACCTTCCAGGTAGCCAGCAGCTCCAGCCGCT 1375
Qy AlaAlaProGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
Dh 1376 RCCCCCTCGGGAGACAGCGCGTGGCGGCTCCCTCCCTCGACCTGTCCAGAGTACC 1435
Qy SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGlnProProPro 180
Dh 1436 AGCAAGAGCGCCAGGTGTCCAGCTTGTGGGAGCAAGAGAGAGAGCCGCCCG 1495
Qy AlaArgSerGlySerGlyGlySerAlaGlyGluProGlnGluGluArgSerGlnGln 200
Dh 1496 GCGAGAGTGGCAGCGCGCGGAGCGCCAGAGAGCCACAGAGAGAGAGAGCCACAG 1555
Qy GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAspAspGlyArgPhe 220
Dh 1556 CAGGATGATATCGAAGAGCTGGAGACCAAGGCGTGGGAATGTCTAACGATGGCGCTTT 1615
Qy LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
Dh 1616 CTCAAGTTTACATCGAAATCGGACAGGCTCTTTAAGACGGTCTCAAAAGGCTCGAC 1675
Qy ThrGluThrThrValGluValAlaTyrPheGluLeuGlnAspArgLysLeuThrLysSer 260
Dh 1676 ACTGAACCAACCGGAGAGTGGCTGGTGTGACCTGACAGATCGAAATTAACAAGTCTT 1735
Qy GlnArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280
Dh 1736 GAGAGCGAGAGATTAAAGAAAGAGCTGAATGTTAAAGGTCTTCAGCATCCCAATATT 1795
Qy ValArgPheLysAspSerTyrPgluSerThrValLysGlyLysGlyIleValLeuVal 300
Dh 1796 GTTAGATTATATGATCTCTGGGAATCCACAGTAAAGAAAGAGTCAATGTGTTGTG 1855
Qy ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
Dh 1856 ACTGAACCTTATGACGTCTGGAACACTTAAACGATCTGAAGAGTTTAAAGTATAG 1915
Qy IleLysValLeuArgSerTyrPheArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
Dh 1916 ATCAAACTTCTAAGAAAGCTGGTGGCTCAAGATCTTAAAGGTCTTCAGTTTCTTCACT 1975
Qy ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Dh 1976 CGAAGCTCACCTTATCATTCACCGGATCTTAAATGTAACAACATCTTATCACCGGCCCT 2035
Qy ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgLysSerPheAla 380
Dh 2036 ACTGGCTCACTCAAGATTTGAGACCTGGTCTGGCAACCTTGAAGCGGGCTTCTTTGCC 2095
Qy LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGlnLysTyrAsp 400
Dh 2096 AAGAGTGTGATAGTACCCAGAGTTCATGAGCCCTCAGATGTATGAGAGAAATATGAT 2155
Qy GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
Dh 2156 GAATTCGTTACGTTATATGCTTTTGGGATGTGATGCTTAAATGCTTAACTCGAATAT 2215
Qy ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
Dh 2216 CTTTACTCGGAGAGGCCAAATGCTGCRCAATCTACCGTGGCTGACCAAGTGGGATAG 2275
Qy ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
Dh 2276 CCGAGCGATTTTGACAAAGTAGCAATTCGAAAGTGAAGAAATTTATTTGAAGATGATA 2335
Qy ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
Dh 2336 CGACAAACAAAGATGAAAGATATTCATCAAGAGACCTTTTGAACATGCGCTTATTCAA 2395

Qy 481 GluGluThrGlyValArgValGluLeuAlaGluLysAspGlyGluLysIleAlaIle 500
Dh 2396 GAGAAACAGAGAGTACGGGTAGAAATTCAGAAAGAGAGATGAGAGAAAAATATGCCATA 2455
Qy LysLeuTrpLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
Dh 2456 AAATATGCTTACGATATTAAGATATTAAGAAATTAAGGAAATTAACAAGATATATCA 2515
Qy AlaIleGluPheSerPheAspLeuGluArgAspValProGluMetValAlaGlnGluMet 540
Dh 2516 GCTATTGAGTTTCTTTTGTATTAAGAGAGATGTCCCAAGATGTGTGACAGAAATG 2575
Qy ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Dh 2576 GTNAGCTGGGATATGCTGTGAGGTGATCACAGACCTGGCTTAAGCTATCAAGAGC 2635
Qy ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGlnGln 580
Dh 2636 AGAGTATCATTAATTAAGAGAAACAGAGCAGCGGAGTTGGTACGAGAGAGCAAGAA 2695
Qy LysLysLysGluGluGluSerSerLeuLysGlnGlnValGluGlnSerAlaSerGln 600
Dh 2696 AAAAAAAGCAGAGAGAGACAGCTCTCAACAGAGGTAGAACATCCAGTGTCCAG 2755
Qy ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Dh 2756 ACAGAGATCAAGAGCTCCCTTCTGTCAGACCGGCTACTACTCTTCTTCACTTCA 2815
Qy AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
Dh 2816 GCTTCAAGTTTCTACACAGTAGAACTGAAGAACTTAGAGCAGATCAACATCAACACTA 2875
Qy GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Dh 2876 CAGTACAGAGAACCGATATATCTGTATCTGATGAGGACGTTGACAGTGTGACAGGA 2935
Qy SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Dh 2936 TCCTGTCTTTCAYAGATCTCAGTACAGAGCAGCCACAGCAGTTCATATGTGCTCCA 2995
Qy HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
Dh 2996 CATGAACAGGACATTTCTACAGGACAGTCCAGGGGATATACCTTACTTACTGACCAAGCA 3055
Qy GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Dh 3056 CAGTCTCAGCCCATGGGTATATCCACCTCAAGTGTGACAGAGGGGAGAGCCAGGAT 3115
Qy GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Dh 3116 CAGCCATCTCAGTACGCTTAAAGGGGTTTCACTTCCCAACCATCAACATCTCTCAG 3175
Qy GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Dh 3176 CAGCAGCAGGAGATATACAGCAGAGCCCTCTCAACAGCAGAGTATTCACCTTCA 3235
Qy GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Dh 3236 CAGACATCAACCTCCAGTGGGCGCCTACTACGACAGCAGGATGAGTARCTCAAGCTCA 3295
Qy GlnValLeuProGlnAlaSerAlaGlyLysGln----- 791
Dh 3296 CAAGTCTTGCTCAAGTATAGCTGGAAGAACAGCTTCCAGTTCCAGCCAGTACCACT 3355
Qy ----- 791
Dh 3356 ATCCAAAGCGAAGCTCAGATCCAGTTGCGACACAAACCTCGGTTGTTCCAGTCACTCT 3415
Qy ----- 791
Dh 3416 GGTGCTCATTTCTTCCAGTGGAGACAGCGGCTCCCTTACCTCTTGTCTCAGTACCT 3475


```
QY 791 ----- 791
Db 3476 GTCTCTARATTCCTCATATCACTCTCATGTCTCTACGGCTCAGACAGAGTTTCTCATCC 3535
QY 791 ----- 791
Db 3536 CTTCCTCATCAATGAGAGCTGAGTACTGAGCCTCTGCTCAGTGTGGCTTCACTGCT 3595
QY 791 ----- 791
Db 3596 ACAACAGCTGCGATCCCGGGGGATCACTGTGTCTCTAGTCTGAGCTTCCAACTCTCTG 3655
QY 791 ----- 791
Db 3656 CAGCCTGTGACTGAGTCCAGTCCAGTTCACCCACAGCTCTCAACAACGAGTTTCCAG 3715
QY 791 ----- 791
Db 3716 TCCATGGGAAATACCAAGTAACTTGGACAAAGCTGTGAGTTCCACTTCTCTGGAGAT 3775
QY 791 ----- 791
Db 3776 GTTCTGTACAGAGGCTTCCACCTCGACTGCGCACACAGTACCCAGAGATTCAAAATAT 3835
QY 791 ----- 791
Db 3836 GCTCCTCTTCCAAAGTGGCTTCTGTTCGATTCATTTCTACAGTCTATCCCTCCCATG 3895
QY 791 ----- 791
Db 3896 CCGACAGAGTACTGGTACACCTGGTACTTTCCACAGTGTGACAGCTTATGTGAA 3955
QY 791 ----- 791
Db 3956 TCAATCTTTTACTTCTCTAGTGGTGTGTAGAGACAGGTTCAAGTGTCCACCCACGA 4015
QY 792 ----- SerThrgln 794
Db 4016 GGGAGTTAGACAGAGCCCACTACATCTCCACAGACAGGTTTGGAGAGTACTCAG 4075
QY 795 GlyAlaSerGlnValAlaProAlaGlnProValAlaValAlaGlnProGlnAlaThrgln 814
Db 4076 GGAATCTCTCAGGTTGCTCTGACAGAGCCAGTTGACAGACAGCAAGTCAAGTCAAG 4135
QY 815 ProThrThrLeuAlaSerSerValAlaSerAlaHisSerAlaValAlaSerGlnMetSer 834
Db 4136 CCGACCACTTTGGCTCTCTCTGTAGACAGTGCACATTCAGATGTGCTTCAAGTATGAGT 4195
QY 835 AspGlnAsnGlnAsnValProSerSerSerGlnArgHisGlnGlyArgThrThrLysArg 854
Db 4196 GATGCAATGAGAGAGCTCCCATCTTCCAGTGAAGCATGAGAGAGCAATACAAACGG 4255
QY 855 HisArgArgLysSerValArgSerArgSerArgHisGlnLysThrSerArgProLysLeu 874
Db 4256 CATTAACGAAATCTGTAGAGAGTCTCTGCACATGAAACCTTCAACGCCCAAAATTA 4315
QY 875 ArgGlnLeuAsnValSerAsnLysGlyAspArgValAlaGlnCysGlnLeuGlnThHis 894
Db 4316 AGAATTTGATGATGTTCAATTAAGAGACCCGAGTATGATGATGATGATGATGATGAT 4375
QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlnAspAsnProGlnLysLeuAla 914
Db 4376 AATAGGAAATGGTACTTCAATTTGACTTACATGATGATGATGATGATGATGATGATGAT 4435
QY 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGlnArgGlnSerPheValAspGln 934
Db 4436 ACAATTTGTGGACCAATGACTTATTTCTGCAATACAGAGAGTGTGATTTGTGATGATGAT 4495
QY 935 ValArgGlnIleIleGlnLysAlaAspGlnMetLeuSerGlnAspValSerValGlnPro 954
Db 4496 GTGCGAATAATTATTTGAAAAAGCTGATGAAATGCTCAGTGAAGTGTGATGATGATGAT 4555
QY 955 GlnGlnLysAspGlnGlnLysLeuSerLeuGlnGlnLysAspAspTyrGlnPheSerGlnSer 974

Db 4556 GAGGATATCAGGAGATTGGAGAGCTTCAAGAGAAAGATATGATGCTTTTCAAGTCT 4615
QY 975 GlnLysLeuGlnGlnGlnLysPheLysGlnProIleProAlaSerSerMetProGlnIle 994
Db 4616 CAGAAATTTGAGAGAGAGTTCAAACCAATTCCTGCTCTTCCATGACAGCAACAATA 4675
QY 995 GlyIleProHisSerSerLeuThrGlnValHisSerLysGlnArgArgPheIleVal 1014
Db 4676 GGCATTCCTACAGATTTTAACTCAAGTGTGATGATGATGATGATGATGATGATGATGAT 4735
QY 1015 SerProValProGlnSerArgLeuArgLysSerLysValPheProSerGlnIleThrAsp 1034
Db 4736 AGTCTGTGCAAGAAACCATTTACAGAAATCAAAATTTTCCCAAGTGAATAACAGAT 4795
QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSer 1054
Db 4796 ACAATTCCTGCTCTACAGCTCAGAGCCCTGGAATGAATGTGCTTCACTGCAATCATCC 4855
QY 1055 LeuSerLeuGlnGlnAlaPheSerGlnLeuArgArgAlaGlnMetThrGlnLysProAsn 1074
Db 4856 CTTAGTCTACACAGGCTTTTCTGAACTTGAAGCTTGAAGCTCCAAATGACAGAGAGCCAA 4915
QY 1075 ThrAlaProHisSerHisSerHisSerThrGlyProThrPheProValValProProPheLeu 1094
Db 4916 ACAGCACTCCAAACTTATGATACAGAGACCAACATTTCCAGTATGATACCTCTTCTTA 4975
QY 1095 SerSerIleAlaGlnValProThrThrAlaAlaAlaThrAlaProValProAlaThSer 1114
Db 4976 AGTAGCATTCCTGGAGTCCCAACACACAGACAGCCACACACAGCCCTGCAACAGC 5035
QY 1115 SerProHisSerHisSerHisSerHisSerValIleGlnSerGlnValThrValProThGln 1134
Db 5036 AGCCTCTCTAATGATCATTTTCCATCATGATGATGATGATGATGATGATGATGATGAT 5095
QY 1135 GlnGlnLysLeuAlaGlnValAlaThrSerThrGlyValValThrSerGlnLysProIle 1154
Db 5096 GAGGAGATTCTGGAGTTCACACAGACAGAGTGTGTAATTCAGTGTGCTCCCATTA 5155
QY 1155 ProProValSerGlnSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
Db 5156 CCACCTGTGTGATGATACACAGTACATTTCCAGGCTATTTCAAGTATCAACAAATACGCA 5215
QY 1175 ValValSerIleSerThrThrThrSerProSerLeuGlnValProThrSerThrSerGlnIle 1194
Db 5216 GTTGTCTCATATCTACTATCATCTCCCTCACTCAATGCTCCCACTCACTCACTCACTCA 5275
QY 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAlaSerAla 1214
Db 5276 GTTGTCTCTAGTACAGACATGATCTCTTCAAGTAAAGTTCAGACATTCAGACATTCAG 5335
QY 1215 GlyGlnSerThrAlaThrProGlnProLysProProAlaValAlaSerGlnGlnAla 1234
Db 5336 GGGGGGAGTACTGATCCCAAGTCTTCAAGCTTCAAGTCTTCAAGTCTTCAAGTCTTCA 5395
QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
Db 5396 GCGACGACTACTGTGGAGCCACATTAACATCACTTCTTCCACCACTTCACTTCCCAAGC 5455
QY 1255 ThrAlaSerGlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGln 1274
Db 5456 ACAGCTTCAACAGCTGTGATCACTTCAAGTACAGACATGATCTTCACTTCACTTCACT 5515
QY 1275 ThrValValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlnLeu 1294
Db 5516 ACCGTGTATGTGGACGACACTCATGATPAAGCACTCATGACATGACATGACATGACATG 5575
QY 1295 AlaPheSerLeuSerAlaProSerSerSerSerSerProGlnAlaGlnValSerSerTyr 1314
Db 5576 GCTTTCCTCTCTGACCATCTTCTCTTCTCTCTCTGAGAGAGAGTGTAGTAT 5635
QY 1315 IleSerGlnProGlnGlnLysLeuHisProLeuValIleProSerValIleAlaSerThrPro 1334
```

Db 5636 ATTCTCAGCTGGTGGGCTGCATCTTGGTCATTCATCAGTATAGCTTACTCT 5695
QY 1335 ILeuProGlnAlaAlaGlyProThSerThrProLeuLeuProGlnValProSerIle 1354
Db 5696 ATTCTCCCAAGAGAGAGAGACTTCTACCTTTATTAACCCAGATCCTAGATC 5755
QY 1355 ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer 1374
Db 5756 CCACCTTGGTACAGCTGTTGGCCAAATGTGCTGCTGTACAGAGACACTAATTCATAGT 5815
QY 1375 GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValAspSer 1394
Db 5816 CAGCTCAACACAGCTTGTCTCCCAACAGCCCACTACTCTTCTGAAAGTATCT 5875
QY 1395 AspThrGlnProLysAlaProGlyIleAspAspIleLeuThrLeuGlnGlnLysLeuArg 1414
Db 5876 GATACACACACCAAGCTCTCGAATGATGACATAAAGACTCAGAAAGAAAGCTCGG 5935
QY 1415 SerLeuPheSerGlnHisSerSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr 1434
Db 5936 TCTCTGTTCCAGTGAACACAGCTCACTGAGCTCAGATGCTCTGTCTCAGTGAAGACC 5995
QY 1435 SerLeuValIleGlnSerThrValThrProGlyIleProThrThrAlaValAlaProSer 1454
Db 5996 TCACTAGTCAATAGAGAGCACTGTCAACACAGCATCCCACTACTGCTGTGACCAAGC 6055
QY 1455 LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr 1474
Db 6056 AAACCTCTGACTTCAACACAGACTTGTCTTACCAACCAACTTACCACTAGAGACA 6115
QY 1475 ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr 1494
Db 6116 GTTGCTTGGCAGTTACACAGAGTGTACACCTGGGCAAGTTCTTACCCAGTCAGACACT 6175
QY 1495 ThrThSerGlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaPro 1514
Db 6176 ACTACATCAGAGACTGAAGACTGGAACCTGCTCTCCAGGCACTTACTTAAGAGCTCG 6235
QY 1515 ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro 1534
Db 6236 GTGCTCCAGTGGGTCTGAACCTTCAGAGGTACTTACCCAGGAGAGAGCTGCCACT 6295
QY 1535 PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu 1554
Db 6296 TTTCAGAGACTTCTCTTAACCCAGTCCAGCAACTCTAGAGATCTTATGCTCAATTG 6355
QY 1555 ArgArgThrLeuSerProGlu**IleThrValThSerAlaValGlyProValSerMet 1574
Db 6356 AGAAGAACTTATGTCAGAGATKATCAAGTACTTCTGGGTGTGTCTGTCTCATG 6415
QY 1575 AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal 1594
Db 6416 GCGGCTCYAAGACATTCACRMAAGAGAGACACACTCAGAAAGGTGTCTTCTCAAGTC 6475
QY 1595 LysGlnGlyProValLeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPhe 1614
Db 6476 AAAAGAGGCTCTCTTACCACTAGTTGAGAGCTGTGTGTTTAAAGTGGACATTT 6535
QY 1615 GlnValSerValAlaAlaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAla 1634
Db 6536 CAGGTTTCTGTGACAGACAGCGTCCAGAAAGAGGATAAATAAGTACAGAAAGTGA 6595
QY 1635 LysSerValHisPheGlnSerSerThrSerGlnSerSerValLeuSerSerSerPro 1654
Db 6596 AAGTCTGTTTCAATTCAGACCTCAGAGTCTCTCATGCTTCAAGTATGATGCTCA 6655
QY 1655 GlnSerThrLeuValLysProGlnProAsnGlyIleThrIleProGlyIleSerSerAsp 1674
Db 6656 GAGAGTACTTGTGTAACCAAGCCGATGACATTAACATCCCTGATATCTCTTCAAGT 6715
QY 1675 ValProGlnSerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlnPro 1694
Db 6716 GTGCAAGAGAGTCCCAAAACTACTGCTCAGAGGCAAAAGTCAAGACCTGGGAGCT 6775

QY 1695 ThrLysValGlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerVal 1714
Db 6776 ACCAAGTGTGACGTTTTCAGGTACCACTTACAGCAACAAAGTGGGTGTTCTGTGA 6835
QY 1715 SerLysThrGlnAspLysIleThrAspThrLysLysGlnGlyProValAlaSerProPro 1734
Db 6836 TCAAAACTAGAGCAAGATCACTGACAAAGAAAGAAAGAGACAGTGGCATCTCTCT 6895
QY 1735 PheMetLysLeuGlnGlnAlaValLeuProAlaValIleProLysLysGlnLysProGln 1754
Db 6896 TTTATGATTTGGAACAGCTGTCTTCTCTGCTGTATACAAAGAAAGAAAGAAAGCTGAA 6955
QY 1755 LeuSerGlnProSerHisLysLeuAsnGlyProSerSerAspProGlnAlaAlaPheLeuSer 1774
Db 6956 CTGTCAAGCTTCACTCACTTAATGGGCGCTTCTGTACCCGAGGCGGCTTTTATAGT 7015
QY 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLysSer 1794
Db 7016 AGGATGTGATGATGTTCCGATGATCCACACTCGCCCATCAGCTGAGCTCAAAAGC 7075
QY 1795 LeuProSerGlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerLysMetSer 1814
Db 7076 CTTCCTAGCCAGATCTAAGTCAAAAGCTTATGATTAATTCATTAATCTCTTACATGAT 7135
QY 1815 SerAspAsnGlnSerAspIleGlnAspGlnAspLeuLysLeuGlnLeuArgLeuArg 1834
Db 7136 AGCGACATATAGTCAATATCGAAGATGAGACTTAAAGTTAAAGTGGCCACACACTACGA 7195
QY 1835 AspLysHisLeuLysGlnIleGlnAspLeuGlnSerAspGlnLysHisGlnIleGlnSer 1854
Db 7196 GATTAACATCTCAAAGATTCAGACCTGACAGAGTCCGACAGATGAATTAATCT 7255
QY 1855 LeuThrThrLysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeu 1874
Db 7256 TTGTATACCAAACTGGGCAAGGTGCCCTGTGTATTATTCCTCCAGCTGCTCCCTT 7315
QY 1875 SerGlyArgArgArgProThrLysSerLysGlySerLysSerArgSerSerSer 1894
Db 7316 TCAAGAGAAAGACAGAGACCTTAACAAAGGACGAACAACTTGTGCAACAGTTC 7375
QY 1895 LeuGlyAsnLysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
Db 7376 TTGGGGAATAAAGGCCCAAGCTTTCAGGTAACTGTCTGTCTCAGAGTGCAGCTTACGTC 7435
QY 1915 LeuHisProGlnGlnThrLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGln 1934
Db 7436 TTGCAACCCCAAGACAGACCTCCACCTCTGCAACATCCAGAGTCCGGGACGAATCAG 7495
QY 1935 LeuLeuGlnProLeuLysProSerProSerSerAspAsnLeuThrSerAlaPheThrSer 1954
Db 7496 CTGTTACAGCCCTTAAAGCATCTCCCTTCACTGACCAACCTTATTCACCTTCAACACT 7555
QY 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsn 1974
Db 7556 GATGGTGCATTTCACTACCAAGCCTTCTGTCTCAGGTCAAGAGAACAGACGACAAAC 7615
QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSer 1994
Db 7616 ACTGTTGGGGCAAGTAAACAGCAAGCCGCGCAAGCTCAGCTCTGCTCATGACGTCC 7675
QY 1995 SerArgLysGlyThrPheThrAspAspLeuHisLysLeuValAspAsnTrpAlaArgAsp 2014
Db 7676 AGCAGAAAGGACATTTCAACAGATACCTTGCACAACTGTGTGACAAATTTGGGCCGAGAT 7735
QY 2015 AlaMetAsnLeuSerGlyArgArgLysSerLysGlyHisMetAsnTrpGlyGlyProGly 2034
Db 7736 GCCATGAATCTCTCAGCAGAGAGAGAGAAAGGACATGAATTAAGAGGCGCTTGA 7795
QY 2035 MetAlaArgLysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGly 2054
Db 7796 ATGCAAGAGAGTCTGCACTGGGCACTGTGACATCTCATATGACCTCAAGCTGAGT 7855

Db	2336	CGCAAAACAAAGATGAAAGATATTCATCAAAAGACCTTTTGAAACCATGCTTTTCCAA	2399
QY	481	GIUGLThGlyValArgValGluLeuAlaGluGluAspGlyGluIleIleIle	500
Db	2396	GAGGAACACAGAGATCCGGGTAGATTTAGCAGAAAGAAAGATGAGAAAAATATACATA	2455
QY	501	LysLeuTriPLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspGln	520
Db	2456	AAATTATGCTACGATTTGAAGATTTAAAGAAATTTAAAGGAAATATCAAAAGTATATGA	2515
QY	521	AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnIleMet	540
Db	2516	GCATTGAGATTTTCTTTTGATTGATTAGAGAGAGATGTCCAGAGAGATGTTGCCAAGAAATG	2575
QY	541	ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp	560
Db	2576	GTAAGACTCTGGGTATGTCTGTAAAGGTGTATCAAAAGCCATGCTTAAAGCTATCAAAAC	2635
QY	561	ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGlnGlu	580
Db	2636	AGAGTATCATTTATATTAAAGGAAAGAGAGCAGCGGACGTTGTGTCGGAGAGACAGAA	2695
QY	581	LysLysLysGlnGluGluSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln	600
Db	2696	AAAAAAAAGCAGAAAGAGAGCGAGTCTCAACACAGCAGTAAAGAACATCCAGTCTCCAG	2755
QY	601	ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer	620
Db	2756	ACGGATATCAAGACAGCTCCCTTCTCTAGCACCGGCATACCTACTGCTTTCACACTTCA	2815
QY	621	AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu	640
Db	2816	GCTTCAGTTTCTACACAGATGAACTGTAAGACCTGAGGAGACATCAACATCAACACTA	2875
QY	641	GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly	660
Db	2876	CAGTACCAAGCAACCCAGATATCTGCTTTACTGATGGACGCTTGACAGTGTCAAGGGA	2935
QY	661	SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln	680
Db	2936	TCCTCTGCTTCAVAGATCTCGAGTGAAGCAGCAACGRCAGTTTCATATAGTTTCCAA	2995
QY	681	HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla	700
Db	2996	CATGAACAGGCACATTTACAGGACAGACAGTCCCAAGGCAATTAACCTTACTGTCCAAACA	3055
QY	701	GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly	720
Db	3056	CAGTCTCAGCCCAAGGGGTATATCCACCTCCAAAGTGTG-----	3094
QY	721	GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln	740
Db	3094	-----	3094
QY	741	GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer	760
Db	3095	-----CAGGGAATACAGCAGACAGCCCTCTCTCAACAGACAGTATTAACCTTCA	3148
QY	761	GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro	780
Db	3149	CAGACATCACTCCAGTGAAGGCCACTACTGCAACAGCAGTGAATCARCCTCAAGCTCCA	3208
QY	781	GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla	800
Db	3209	CAAGTCTTGCTCAAGTATCACTGCGAAAAACAGATCTCAGGGAGTCTTCAGGTGTCT	3268
QY	801	ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer	820
Db	3269	CCTGAGAGCCAGTTGACGTAGCACAACACCCAACTACCAAGCCGACCACTTTGCGCTTCC	3328
QY	821	SerValAspSerAlaHisSerAspValAlaAspGlyMetSerAspGlyAsnGluAsnVal	840
Db	3329	TCTGTAGACATGACATTCAGATGTTGCTTCAGATATGATGATGGCATTTGAACGTC	3388

QY	841	ProSerSerSerGlyValArgHisGluGluValArgThrThrIlySaRHisTyrArgIysSerVal	860
Db	3389	CCATCTTCCAGGGAAGGACATGAAAGGAACCTACAAAACGGCATTTACGAAAATCTGTA	3448
QY	861	ArgSerArgSerSerArgHisGluIlySerThrSerArgProIlyLeuArgIleLeuAsnValSer	880
Db	3449	AGGAGTCCGCTCTCCGACATGAAAACTTCAGCCCCAAAATTAGAAATTTGAAATTTTCA	3508
QY	881	AanIySGIYASpARgValValGluCysGlnLeuGluThrHisSaSaArgIysMetValThr	900
Db	3509	AATAAAGGAGACCCGAGTGTAGAAATGTCAATTAGAGACTCATPATATAGAAAAATGTTTACA	3568
QY	901	RhelysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn	920
Db	3569	TTCAAAATTTGACCTAGATGATGTGACAAACCCAGAGGATAGCAACAATTATGTGTGAACAT	3628
QY	921	AspPheIleLeuAlaIleGluArgGluSerPheValAspGluValArgGluIleIleGlu	940
Db	3629	GACCTTATTTCTAGCAATAGAGAGAGAGTCGTTTGTGGATCTAAGTCCGAAAATTTATGAA	3688
QY	941	LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGluYAspGlnGlyLeu	960
Db	3689	AAAGCTGATGAATCTCAGTAGGAGATCTAGTGTGGAACACAGAGGGTATCAGGATTTG	3748
QY	961	GluSerLeuGlnGlyIlyAspAspArgArgIlyPheSerArgIlyLeuLeuGluGlyGlu	980
Db	3749	GAGAGTCTACAGAAAGAAAGATGACTATGCTTTTCAAGTCTTCGAAATTTGAAAGGAG	3808
QY	981	PheIySGInProIleProIleAspSerSerMetProGlnGlnIleGlyIleProThrSerSer	1000
Db	3809	TTCAACACCAACAAATTCCTGCGTCTTCCATGCACAGCAAAATAGGCATTCTCAACAGTTCT	3868
QY	1001	LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGluSer	1020
Db	3869	TTAACTCAAGTTGTTCATCTGCGGGAAGGGGTTTATAGAGATCTCTGTGCCAAGAAC	3928
QY	1021	ArgLeuArgGluSerIlyValPheProSerGluIleThrAspThrValAlaIleSerThr	1040
Db	3929	CGATTAGAGAAATCAAAAGTTTCTCCAGTAAATACAGATACAGTTGTCCTCTACA	3988
QY	1041	AlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeuGlnGlnAla	1060
Db	3989	GCTCAGACCCCTGGAAATGAATCTTCTCACCCTTCATCATCCCTTAGCTTACAACAAGCC	4048
QY	1061	PheSerGluLeuAspArgAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe	1080
Db	4049	TTTTCTGAACTTAAGCTGCCCAATAGACAAAGAACCCAAANACAGACCTCCAAACTTT	4108
QY	1081	SerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGlyVal	1100
Db	4109	AGTCATACAGGACCAACATTTCCAGTAGTACCTCTTTCTTAAGTACATTTGCTGAGTCC	4168
QY	1101	ProThrThrAlaAlaAlaThrAlaProValProAlaHisSerSerProProAsnAspIle	1120
Db	4169	CCAACTVACAGAGACCAACAGCACCGACTCTTGCAACAAGCCCTCTTAATGACATT	4228
QY	1121	SerThrSerValIleGlnSerGluValThrValProThrGluGlnGlyIleAlaGlyVal	1140
Db	4229	TTCACATCAGTAATTCAGTCTGAGGTTACAGTGGCCCACTAAGAGGGGATTTGCTGAGTT	4288
QY	1141	AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer	1160
Db	4289	GCCACACAGACAGGTGTGTAACTTCAGGTGTCTCCCACTACCACTGTGTCTGATCA	4348
QY	1161	ProValIleuSerSerValIleSerSerIleThrIleProAlaValIleSerIleSerThr	1180
Db	4349	CCAGTACTTTCACGCGTAGTTTCAAGTATACAAATACCTCAGTGTCTCAAAATCTACT	4408
QY	1181	ThrSerProSerLeuGlnValProThrSerThrSerGluIleValIleSerSerThrAla	1200
Db	4409	ACATCCCCGTCACCTCAAGTCCCAATCCCAATCTGAGATCGTTGTCTTATATACAGA	4468

Db 6629 CTCACCCCTCTGGGACATCCAGAGCCGGGACGAATACGCTTTACAGCCCTTTAAG 6688
 Qy 1941 ProSerProSerSerAspAsnLeuTyrSerAlaPheThrSerAspGlyAlaIleSerVal 1960
 Db 6689 CCATCTCCCTCCAGAGACACCTCTATTACAGCTTACACAGTATGGGCAATTTCAGTA 6748
 Qy 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
 Db 6749 CCAAGCCCTTCTGCTCCAGGTCAGAGAACCCAGACACAAACATGTTGGGGCAACAGTG 6808
 Qy 1981 AsnSerGlnAlaIleAlaGlnProProAlaMetThrSerSerArgIleValThrPhe 2000
 Db 6809 AACGCGCAAGCCGCCCAAGCTCAGCTCTCCAGACGTCACAGCAGAGAGGACATTC 6868
 Qy 2001 ThrAspAspLeuHisIleValLeuValAspAsnTrpAlaArgAspAlaMetAsnLeuSerGly 2020
 Db 6869 ACAGATGACTTGACCAAGTTGGTAGACATTGGGCCCGGAGATGCCATGAACTCTTCAGGC 6928
 Qy 2021 ArgArgGlySerIleGlyHisMetAsnTrpGlnGlyProGlyMetAlaArgIlePheSer 2040
 Db 6929 AGGAGAGAGAACCAAGGACATGATATTAGAGGGCCCTGGAAATGGCAGAGAAATTCCT 6988
 Qy 2041 AlaProGlyGlnLeuGlyIleSerMetThrSerAsnLeuGlyIleSerAlaProIleSer 2060
 Db 6989 GCACCTGGGCACTGTGATCTCCATGACCTCGAACTGGGTGGCTCTGCCCTCATCTCT 7048
 Qy 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrIleSerMetCysProProGlnGlnTyr 2080
 Db 7049 GCAGCATCAGCTACTCTCTAGTCACTTACCAAGCTATAGTCCGCCCCACAGCAGATAT 7108
 Qy 2081 GlyPheProAlaThrProPheGlyAlaGlnTyrSerGlyThrGlyIleProAlaProGln 2100
 Db 7109 GGGCTTTCAGGTACCCCAATTTGGGCTCAATGGAGTGGAGCGGTGGCCACAGCACAG 7168
 Qy 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
 Db 7169 CCACCTGGGCACTTCCAACTGTGGGAACTGCTCTTGGCAATTTCAACATCAGCAAT 7228
 Qy 2121 LeuGlnIleSerIleSerAsnProProGlySerAsnLeuArgThrThr 2136
 Db 7229 TTGGAGAAATCCATCAGCAACCCCGAGGCTCCCACTGGGGAGCACT 7276
 RESULT 9
 US-10-767-471-452
 ; Sequence 452, Application US/10767471
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001505
 ; CURRENT APPLICATION NUMBER: US/10/767,471
 ; CURRENT FILING DATE: 2004-01-30
 ; NUMBER OF SEQ ID NOS: 50231
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 452
 ; LENGTH: 7853
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(7853)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
 US-10-767-471-452

Alignment Scores:

Pred. No.: 0 Length: 7853
 Score: 10567.50 Matches: 2096
 Percent Similarity: 98.13% Conservative: 0
 Best Local Similarity: 98.13% Mismatches: 11
 Query Match: 97.74% Indels: 29
 DB: 7 Gaps: 1

US-10-010-720-14 (1-2136) * US-10-767-471-452 (1-7853)

Qy 1 MetSerGlyGlyAlaIleAlaGlnIleSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 956 ATGCTGGCGCGCGCGAGACAGACAGACAGCACTCCGGTTCCTCTCTCCGCG 1015
 Qy 21 ProAlaProAlaProIleValSerGlySerSerSerAspSerValGlyGlyLeuGly 40
 Db 1016 CCGGCTCTCCGCGCGAGAAATGGCTCAGCTCCGATTCCTCCGTGGGGAGAAACTGGGA 1075
 Qy 41 AlaAlaAlaAlaSerAlaValThrGlyArgThrGlnIleTyrArgAlaGlyArgHisThr 60
 Db 1076 GCGGCGCGCGCGAGCTGTGACCGGAGAGACCAAGAGATACAGCCCGCGCGACACT 1135
 Qy 61 MetAspIleAspSerSerArgGlyAlaIleAlaIleThrThrThrThrThrThrThrThr 80
 Db 1136 ATGACACAGACAGACAGCTGGGGCGGGCGGCGCACCACTTCCACCACTTGGACAGCTTCTC 1195
 Qy 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGlnLeuProGlyLeuProLeu 100
 Db 1196 CGCGGAGCGTCACTCTGNGACTCCAAATGCCACTGCACTGGAGCTTCCCGGCTTCTCTT 1255
 Qy 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGlnProHis 120
 Db 1256 TCCCTGCCCGAGCCAGACATCCCGCGCTGCTCCGAGAGTGTCCACCGGAGACCCGAC 1315
 Qy 121 ArgGlnIleThrValThrAlaThrAlaThrSerGlnValAlaGlnIleProProAlaAla 140
 Db 1316 CGGGAAGACACCGTGAAGCCGACCGCACTTCCAGTACCCAGACAGCTTCCAGCGCT 1375
 Qy 141 AlaAlaProGlyGlnAlaValAlaIleGlyProAlaProSerThrValProSerSerThr 160
 Db 1376 RCGGCCCTGGGGAACAGCGCTGCGGGGCTGCGCCCTCGACTGTGCCAGAGATACC 1435
 Qy 161 SerIleAspArgProValSerGlnProSerLeuValGlySerIleGlnIleProProPro 180
 Db 1436 AGCAAAACCGCCAGAGTCCAGCCTTGTGGGAGCAAAAGAGAGAGCGCCCGCG 1495
 Qy 181 AlaArgSerGlySerGlyGlyIleSerAlaIleGlnProGlnIleGlnIleValSerGlnGln 200
 Db 1496 GCGAGAAATGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1555
 Qy 201 GlnAspAspIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 220
 Db 1556 CAGGATGATATCGAAGAGCTGAGAGACCAAGCGCTGGAAATGCTAACATGGCGCTTT 1615
 Qy 221 LeuIlePheAspIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 240
 Db 1616 CTCAGATTGACATCGAAATCGGACAGAGCTCTTTAAGCGGTCTTACAAAGGCTTGAC 1675
 Qy 241 ThrGlnThrThrValGluValAlaIleTPCysGlnLeuGlnAspArgIleThrIleSer 260
 Db 1676 ACTGAACACCGCTGGAAGCTCGCTGTGTGAATCGAGATGGAATTAACAACTCT 1735
 Qy 261 GluArgGlnArgPheIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 280
 Db 1736 GAGAGCAGAGATTTAAGAAAGCTGAAATGTTAAAGCTTCCAGCATCCCAATATT 1795
 Qy 281 ValArgPheIleIleAspIleIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 300
 Db 1796 GTTAGATTATTAATTCCTGGGAATCCACAGTAAAGAAAGAAAGTGCATGTTTGGTG 1855
 Qy 301 ThrGlnLeuMetThrSerGlyThrLeuIleThrIleThrIleThrIleThrIleThrIle 320
 Db 1856 ACTGAACCTATGACGCTGGAACACTTAAACCTATCTGAAAAGGTTTAAAGATGACAG 1915
 Qy 321 IleIleValIleuArgSerTrpCysArgGlnIleLeuIleGlyLeuGlnPheLeuHisThr 340
 Db 1916 ATCAAAAGTTCTAAGAGAGCTGTCGCTCAGATCTTAAAGTCTTCATCACTACT 1975
 Qy 341 ArgThrProProIleIleIleHisArgAspLeuIleCysAspAsnIlePheIleThrGlyPro 360
 Db 1976 CGAACTCCACTATCAATCCGCGGATCTTAATGTGACAAACATTTATCAGCGGCGCT 2035

QY 361 ThrGlySerValIleGlyIleAspLeuGlyLeuAlaThrLeuLysArgAlaSerPheAla 380
 DB 2036 ACTGGCTCACTCAAGATTGGAGACCTGCTGGCAACCTGAAGCGGGCTTTTGGCC 2095
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluTyrAsp 400
 DB 2096 AAGAGTGTGATTAGTACCCAGAGGTTGATGCGCCCTGAGATGATGAGGAATATGAT 2155
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
 DB 2156 GAATCCCTTACCTTATGCTTTTGGATGTCAGATGTCAGATGTCAGATGTCAGAT 2215
 QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValIle 440
 DB 2216 CTTTACTCGAGTCCAAAATGCTGCGCAGATCTTACCTGCGCGTACCAGTGGGGTAA 2275
 QY 441 ProAlaSerPheAspLysValAlaAlaIleProGluValLysGluIleIleGluGlyCysIle 460
 DB 2276 CCAAGCCAGTTTTCAGAAAGTAGCAATTCCTGAAGTGAAGAAATTTATGAAGATGCATA 2335
 QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuMetAlaPhePheGln 480
 DB 2336 CGACAAAACAAAGATGAAGATATTCATCAAGAACCTTTTGAACATGCTTTTCCAA 2395
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
 DB 2396 GAGGAACACAGAGTACCGGTGATGATTCAGAGAACATGATGAGAAATTAATGCCATA 2455
 QY 501 LysLeuThrLeuArgIleGluAspIleLysLysLeuLysGlyTyrTyrLysAspAsnGlu 520
 DB 2456 AAATTATGCTACGATATTGAAGATATTAGAAATTTAAAGGGAATTAAGAAATATGAA 2515
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 DB 2516 GCTATTAGTATTTCTTTGATTAGAGAGATGTCCTCAGAAAGTGTGCAACAAAGATG 2575
 QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 DB 2576 GTAGAGTCTGGTATGCTCTGGAAGTGCATCAAGAACATGCTTAAAGCTATCAAAAGAC 2635
 QY 561 ArgValSerLeuIleLysArgLysArgGluGluArgGluLeuValArgGluGluGlu 580
 DB 2636 AGAGTATCATTAATTAGAGAGAAACGAGACGCGAGTGTGATCCGAGAGAGAACAA 2695
 QY 581 LysLysLysGluGluGluSerSerLeuLysGlnGlnAlaGluGlnSerSerAlaSerGln 600
 DB 2696 AAAAAAAGCAGAGAGAGAGAGATCTCAACACAGAGTGAACAAATCCAGTGTCCAG 2755
 QY 601 ThrGlyIleLysGluLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 DB 2756 ACGAGATCAAGACGCTCCCTTCTGCTAGCACCGGCACTACTGCTTCAACACTTCA 2815
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGluLeu 640
 DB 2816 GCTTCACATTTCTCAACAGTAGAAGAACCTGAAGAACCTGAGCAGATCAACATCAACAACTA 2875
 QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 DB 2876 CAGTACAGCAACCAAGATATCTGTGTATCTATGAGGAGGTTGACAGTGGTCAAGGGA 2935
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 DB 2936 TCCTCTGCTTCAAGATCTCGAGTGAAGCCCAACAGRCAGTTTCAATNGTCTCCAA 2995
 QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 DB 2996 CATGAACAGGACATTTCTACAGCACAGTCCCAAGGACATATACCTTACTGCTCAAGCA 3055
 QY 701 GluSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnGlnSerGlnGly 720
 DB 3056 CAGTCTAGGCCCATGGGGTATATCCACCTCAAGTGG----- 3094
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740

DB 3094 ----- 3094
 QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
 DB 3095 -----CAGGAATATACGACGACAGACGCCCTCCCAACAGACAGTGCAGTATTCATCTTCA 3148
 QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 DB 3149 CAGACATCAACCTTCAGTGAAGCCACTACTGCAAGCCAGTGAAGTCAACCTTCAAGCTTCA 3208
 QY 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
 DB 3209 CAAGTCTGCTCAAGATACAGCTGGAAGAAACAGAGTACTCAGGAGAGTCTCTCAGGTTGCT 3268
 QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
 DB 3269 CCTGACAGGCCAGTTCGATGACAGCCAGCCCAAGCTAACCCAGCCGACCTTGGCTTCC 3328
 QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
 DB 3329 TCTGTACACGTGCACATTCAGATGTTGCTTCAGATGATGATGATGATGATGATGATGAT 3388
 QY 841 ProSerSerSerGlyArgHisGlnGlyArgThrThrLysArgHisTyrArgLysSerVal 860
 DB 3389 CCATCTTCAGTGAAGGCATGAAGAAAGAACTAACAAACGCGCATTCGAAAATCTGTA 3448
 QY 861 ArgSerArgSerArgHisGlyLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
 DB 3449 AGAGTGTGCTCTGACATGAAAAAACTTACGCCCCAAATTTAAGAAATTTGAATGTTC 3508
 QY 881 AsnLysGlyAspArgValValGluCysGlnLeuGlnThrHisAsnArgLysMetValThr 900
 DB 3509 AATTAAGAGAACCGAGTGAAGATGTCATTTAGATGATGATGATGATGATGATGATGATGAT 3568
 QY 901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn 920
 DB 3569 TTCAAATTTGACTAGTATGCTGACCAACCCGAGAGATGACCAATTAATGATGACAAAT 3628
 QY 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
 DB 3629 GACTTATTTCTCAATGAAGAGAGAGTGTGTTGATCAAGTGCAGAAATTAATTTGA 3668
 QY 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGlyAspGlnGlyLeu 960
 DB 3668 AAAGCTGATGAATGCTCACTGAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 3748
 QY 961 GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGlnGlyGlu 980
 DB 3749 GAGAGTCTACAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3808
 QY 981 PheLysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
 DB 3809 TTCAAAACCAACCAATTCGTGCTTCCATGCTCCACAGCAAAATAGGCATTTCTTCAACAGTCT 3868
 QY 1001 LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGluSer 1020
 DB 3869 TTAACTCAAGTGTGATCTTCTGCGGAGAGCGGTTTATATGATGATGATGATGATGATGATGAT 3928
 QY 1021 ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr 1040
 DB 3929 CGATTACGAACCAAAAGTTTCCCAAGTGAATAACAGATACAGTGTGCTCTTCA 3988
 QY 1041 AlaGlnSerProGlyMetLeuLeuSerHisSerAlaSerSerLeuSerLeuGlnAla 1060
 DB 3989 GCTCAGAGCCCTGGAAGAACTGTCTCAGCTGCAATCACTTCTTACTTCAACAGAGCC 4048
 QY 1061 PheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe 1080
 DB 4049 TTTTCTCAACTTACAGCTGCCAAATGACAGAAAGAACCCAAACAGACCTCCCAACTTT 4108
 QY 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal 1100

Db 4109 AGTCATACAGAGCAACATTTCCAGTACCTCTTCTTAAGTACATTCGAGATC 4168
Qy 1101 ProthThrrAlaAlaAlaThrrAlaProValProAlaThrrSerSerProProAsnAspIle 1120
Db 4169 CCAACVACAGAGAGAGCCACAGACACAGTCCCTGGACACAGAGCCCTCTTAATGACATC 4228
Qy 1121 SerThrrSerValIleGlnSerGluValThrrValProThrrGluGluValIleAlaGlyVal 1140
Db 4229 TCCACATAGTAATTCAGTCTGAGGTACAGTGCCTGACGAAAGGGGATTCGCGAGTT 4288
Qy 1141 AlaThrrSerThrrGlyValValThrrSerGlyGlyLeuProIleProProValSerGluSer 1160
Db 4289 GCCACAGAGCAGAGGTGGTACTTCAGGTGGTCTCCCATACACCTGTGTCTGAATCA 4348
Qy 1161 ProValLeuSerSerValValSerSerIleThrrIleProAlaValIleSerIleSerThr 1180
Db 4349 CCAGTACTTTCACAGTATTCAGATATCACATATCCTGACATGATGCTCATATCTACT 4408
Qy 1181 ThrSerProSerLeuGlnValProThrrSerThrrSerGluIleValIleSerSerThrrAla 1200
Db 4409 ACATCCCGTACTTCAAGTCCCAACATCCACATCAGAMTCGTTGTTCTAGTACAGCA 4468
Qy 1201 LeuThrrProSerValThrrValSerAlaThrrSerAlaSerAlaGlySerThrrAlaThr 1220
Db 4469 CTGATCTTCAGTAACAGTTTCAGCAATTCAGCTCTGCAAGGGGAGTACTGCTACC 4528
Qy 1221 ProGlyProIlyProProAlaValIleSerGlnAlaAlaGlySerThrrThrrValGly 1240
Db 4529 CCAGGTCTTAAGCTCCAGCTGATATCTCAGCAGCAGCAGCAGCAGCAGTCTGCGA 4588
Qy 1241 AlaThrrLeuThrrSerValSerThrrThrrThrrSerPheProSerThrrAlaSerGlnLeuSer 1260
Db 4589 GGCACATTAACATCAGTTTCTACACACATTCATCCCAAGCAGCAGCTCTGACGCTGTC 4648
Qy 1261 IleGlnLeuSerSerSerThrrSerThrrProThrrLeuAlaGluThrrValIleValSerAla 1280
Db 4649 ATTACGTTTACAGACGACTTCTTACTCTCTTACTTACTTAAACCGGTGATGAGCGCA 4708
Qy 1281 HisSerLeuAspIlyThrrSerHisSerSerThrrThrrGlyLeuAlaPheSerLeuSerAla 1300
Db 4709 CACTCATTATATAGACATCTCATAGCAGTACCACTGATTTGGTTTCTCTCTCTGCA 4768
Qy 1301 ProSerSerSerSerSerProGlyAlaGlyValIleSerSerThrrIleSerGlnProGlyGly 1320
Db 4769 CCACTCTCTCTCTCTCTCTGAGCAGCAGAGTGTCTAGTATATTTCTCAGCCTGGGCG 4828
Qy 1321 LeuHisProLeuValIleProSerValIleAlaSerThrrProIleuProGlnAlaAla 1340
Db 4829 CTGCATCTTGGTATTCATCAGATGATGCTTCTATCTCTTATCTTCCCAAGCAGCA 4888
Qy 1341 GlyProThrrSerThrrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
Db 4889 GGACCTACTTCTACACTTATATACCCAGATAGTATCCACCCCTTGGTACAGCT 4948
Qy 1361 ValAlaAsnValProAlaValGlnGlnThrrLeuIleHisSerGlnProGlnProAlaLeu 1380
Db 4949 GTTGCCAAATGTGCTCTGTACAGCAGACACTAATTCATGTACAGCTCAACACAGCTTG 5008
Qy 1381 LeuProAsnGlnProHisThrrHisGlyProGlnValIleAspSerAspThrrGlnProIlyAla 1400
Db 5009 CTTCACACAGCCCATCTACTCTTGTCTGAAAGTATCTGATACACCAACCAAGCT 5068
Qy 1401 ProGlyIleAspAspIleThrrThrrLeuGluGluIlyLeuAspSerLeuPheSerGluHis 1420
Db 5069 CCGGAAATGATGACATTAAGACTTGAAGAAAGAACTGGGTCTGTCTTCACTGAAAC 5128
Qy 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrrSerLeuValIleGluSer 1440
Db 5129 AGCTATCTGAGAGTACAGATGCTCTGCTCACTGAGAGCCTCACTAGCATAGAGAGC 5188
Qy 1441 ThrValThrrProGlyIleProThrrThrrAlaValAlaProSerIlyLeuLeuThrrSerThr 1460
Db 5189 ACTGTACACAGCAGCATCCCACTACTGCTGTGGACCAAGCAAACTCTGACTTCTACC 5248

Qy 1461 ThrSerThrrCysLeuProProThrrAsnLeuProLeuGlyThrrValAlaLeuProValThr 1480
Db 5249 ACAAGTACTGTCTTACACACACCAATTTTACCATAGAAACAGTTCCTTCCGCTTACA 5308
Qy 1481 ProValValThrrProGlyGlnValSerThrrProValSerThrrThrrThrrSerGlyValIly 1500
Db 5309 CCAGTGTACACCTGGGCAAGTTTCTTACCCCACTGACACTATCATCATCAGAGTGA 5368
Qy 1501 ProGlyThrrAlaProSerIlyProProLeuThrrIlyAlaProValLeuProValGlyThr 1520
Db 5369 CTTGAACTGCTCTCTTCAAGCACAACCTTAAGTCAAGCTCGGTGCTGCAAGTGGTACT 5428
Qy 1521 GluLeuProAlaGlyThrrLeuProSerGluGlnLeuProPheProGlyProSerLeu 1540
Db 5429 GAATCTCCACAGGATCTTACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 5488
Qy 1541 ThrGlnSerGlnGlnProLeuGluAspLeuAspAlaGlnLeuArgThrrLeuSerPro 1560
Db 5489 ACCCAGTCCAGCAACCTTAGAGATCTTGAATGCTCAATTGAGAAAGACACTTATGCCA 5548
Qy 1561 Glu***IleThrrValThrrSerAlaValAlaGlyProValSerMetAlaAlaProThrrAlaIle 1580
Db 5549 GAATATATCACAGTACCTTCTGCGGTGCTGCTGTCTCATGCGGCTCYAACAGCAATC 5608
Qy 1581 ThrGluAlaGlyThrrGlnProGlnIlyGlyValSerGlnValIlyGluGluProValLeu 1600
Db 5609 ACRGAAGCAGAGACACAGCTCCAGAGAGGTGTTTCTCAATCAAGAAAGGCCCTGTCTTA 5668
Qy 1601 AlaThrrSerSerGlyAlaGlyValIlePheIlySerGlyArgPheGlnValSerValAlaAla 1620
Db 5669 GCAACTAGTTCAGAGAGTGGTGTGTTTAAAGATGGAGCAATTCAGGTTTCTGTGGCAGCA 5728
Qy 1621 AspGlyAlaGlnIlyGluGlyIlyAsnIlySerGluAspAlaIlySerValHisPheGlu 1640
Db 5729 GACGGTCCCAAGAAAGGATTAATAATGTCAGAAAGTCCAAAGCTGTCTTCAATTTGAA 5788
Qy 1641 SerSerThrrSerGluSerSerValLeuSerSerSerSerProGluSerThrrLeuValIly 1660
Db 5789 TCCAGACCTCAGAGTCTCTCAGTCTTATCAAGTATGTCAGAGATACCTTGGTGA 5848
Qy 1661 ProGluProAsnGlyIleThrrIleProGlyIleSerSerAspValProGluSerAlaHis 1680
Db 5849 CCAGAGCCGAGATGCATTAACATCCTGATCTTCTAGATGTGCAGAGAGTGCCTAC 5968
Qy 1681 IlyThrrThrrAlaSerGluAlaIlySerSerAspThrrGlyGlnProThrrIlyValIlyArgPhe 1700
Db 5909 AAAACTACTGCTCTCAGAGCAGCAAGTCAGACACTGGGCAAGCTTACCAAGTTGACGTTT 5968
Qy 1701 GlnValThrrThrrAlaAsnIlyValIlyArgPheSerValSerIlyThrrGluAspIly 1720
Db 5969 CAGGTACAACTACAGCAACCAAGTGGTCTTCTGTATCAAAAACGTGAGCAAC 6028
Qy 1721 IleThrrAspThrrIlyIlyGluGlyProValAlaSerProPheMetAspLeuGluGln 1740
Db 6029 ATCACTGACACAAAGAAAGAGAGCAGTGCATCTCTCTTTATGATTTGGAAACA 6088
Qy 1741 AlaValLeuProAlaValIleProIlyIlyGluIlyPProGluLeuSerGluProSerHis 1760
Db 6089 GCTGTCTTCTCTGCTGTATACCAAGAAAGAAACCTGAACTGCAGAGCTTCAAT 6148
Qy 1761 LeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspValAspAspGly 1780
Db 6149 CTAAATGGGCGCTCTTCTTACCCCGAGGCGCTTTTAAAGTGGGATGTGATGATG 6208
Qy 1781 SerGlySerProHisSerProHisGlnLeuSerSerIlySerSerLeuProSerGlnAsnLeu 1800
Db 6209 TCCGATAGTCCACATCGGCCCATCATCAGTGAAGTCAAAAGCTTCTTCCAGATCTA 6268
Qy 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerIlyMetSerSerAspAsnGluSerAsp 1820
Db 6269 AGTCAAAAGCTTATGTAATTCATTTAACTCTCTTACATGATGAGCAGCAATAGTCAAT 6328

Db 1676 ACTGAACACCGTGAAGTGCCTGCTGATCGACGATCCAAAAATTACAAAGCTCT 1735
Qy 261 GUAAGGUAAPhlyeGluGluAglMetLeuysGlyLeuGlnHisProAsnIle 280
Db 1736 GAGAGGCGAGATTTAAAGAGAGAGCTGAATGTTAAAGAGCTTCCAGCATCCCAATATT 1795
Qy 281 ValArgPheTyraAspSerTrpGluSerThrValIysGlyLysLysCysIleValLeuVal 300
Db 1796 GTTAGATTATTAATGATCTCTGGGAATCCACAGTAAAGAGAGAGAGATGATGTTTGCTG 1855
Qy 301 ThrGluLeuMetTrsSerGlyThrLeuLysThrTrpLeuLysArgPheLysValMetLys 320
Db 1856 ACTGAACCTTAGAGCTCTGGAACCTTAAGAGATCTGAAGAGAGGTTTAAAGTATGAAG 1915
Qy 321 IleLysValLeuArgSerTrpCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
Db 1916 ATCAAGGTTCTAAGAGCTGCTGCGTCAAGATCCTTAAAGGCTTCCAGTTTCTTCATCT 1975
Qy 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
Db 1976 CGAATCTCACCCTACATTCACCGGATCTTAATGTACAAACATCTTATCACCGGCCCT 2035
Qy 361 ThrGlySerValLysIleGlyAspLeuGlyLeuValThrLeuLysArgAlaSerPheAla 380
Db 2036 ACTGCTCAGTCAGCAAGATTGAGAGACCTCGTCTGGCAACCTTGAAAGCGAGCTTCTTTGCC 2095
Qy 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTrpGluLysTyraSp 400
Db 2096 AAGAGTGTAGTAGTACCCAGAGGTTTCATGCGCCCTGAGATGTATGAGAGAAATATGAT 2155
Qy 401 GluSerValAspValTyraIlePheGlyMetCysMetLeuGluMetAlaThrSerGlyTrp 420
Db 2156 GAATTCGTTACGTTATAGCTTTTGGAGATGTGCAATGCTGATGCTTACATCTGAATAT 2215
Qy 421 ProTyrSerGluCysGlnAsnAlaIleGlnIleTyraArgArgValThrSerGlyValLys 440
Db 2216 CTTACTCGGAGTGCAGAAATGCTGCGACATCTTACCTGCGTGCAGGAGTGGGGTGAAG 2275
Qy 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
Db 2276 CCAGCCAGTTTTCAGCAAGTAGCAATTCCTGAGGTAGAGAAATTTATGAAAGATGATA 2335
Qy 461 ArgGlnAsnLysAspGluAlaGlyTyrSerIleLysAspLeuLysAsnIleAlaPhePheGln 480
Db 2336 CGACAAACAAAGATGAAAGATATTCATCAAAAGACCTTTTGAACATGCTTTTCCAA 2395
Qy 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
Db 2396 GAGGAAACAGAGATGACGGTGAATTAAGCAAGAGAGATGAGAGAGAGAGAGAGAGAGAG 2455
Qy 501 LysLeuTrpLeuArgIleGluAspIleLysLysLeuLysGlyLysTyraLysAspAsnGlu 520
Db 2456 AAATTAATGCTACGATTAAGATATTAAGAAATTTAAAGGAAATTAACAAAGATTAATGA 2515
Qy 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
Db 2516 GCTAATGAGTTTCTTTGATTATTAAGAGAGAGATGCTCCAGAGAGATGTTGACAAAGAAATG 2575
Qy 541 ValGluSerGlyTyraValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
Db 2576 GTTAGATCTGGAGTATGCTGTGAAGGATGATCACAGAGCAATGGCTTAAAGCTATCAAGAGC 2635
Qy 561 ArgValSerLeuIleLysArgLysArgGluGlnArgIleLeuValArgGluGlnGlu 580
Db 2636 AGAGTATCATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2695
Qy 581 LysLysLysGluGlnGluLysSerLysLysGlnGlnValGluGlnSerAlaSerGln 600
Db 2696 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2755
Qy 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrTrs 620
Db 2756 AAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2815

Qy 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGluLeu 640
Db 2816 GCTTCAGTTTCTACAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2875
Qy 641 GlnTrpGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 2876 CAGTACAGAGAGAGAGAGAGATATGTTATCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 2935
Qy 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyraGlySerGln 680
Db 2936 TCCCTGCTCTTCAGAGAGATCTCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2995
Qy 681 HisGlnGlnAlaHisSerThrGlnGlyThrValProGlyHisIleProSerThrValGlnAla 700
Db 2996 CATGAACAGGAGAGATTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3055
Qy 701 GlnSerGlnProHisGlyValTyraProProSerSerValAlaGlnGlyGlnSerGlnGly 720
Db 3056 CAGTCTCAGCCCAATGGGATATATCCACCTCAAGTGTG----- 3094
Qy 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
Db 3094 ----- 3094
Qy 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnIleTyraSerLys 760
Db 3095 -----CAGGAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3148
Qy 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 3149 CAGACATCAACCTCCAGTAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3208
Qy 781 GlnValLeuProGlnAlaSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
Db 3209 CAAGCTGCTGCTCAAGTATCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3268
Qy 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrTrpLeuAlaSer 820
Db 3269 CTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3328
Qy 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
Db 3329 TCTGTAGACAGTCAATTCAGATGTTGCTTCAGATGATGATGAGATGAGAGAGAGAG 3388
Qy 841 ProSerSerSerGlyArgHisGluGlyArgGlnThrLysArgHisTyraArgLysSerVal 860
Db 3389 CCATCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3448
Qy 861 ArgSerArgSerArgHisGluLysThrSerArgProLysLeuArgIleLeuAsnValSer 880
Db 3449 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3508
Qy 881 AsnLysGlyAspArgValAlaGluCysGlnLeuGluThrHisAsnArgLysMetValThr 900
Db 3509 AATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3568
Qy 901 PheLysPheAspLeuAspGlyAspAsnProGluGluIleAlaThrIleMetValAsnAsn 920
Db 3569 TTCAAATTTGACCTTAATGATGATGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3628
Qy 921 AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu 940
Db 3629 GACTTATTTCTACAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3688
Qy 941 LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGlyAspGlnGlyLeu 960
Db 3689 AAGAGTGAATGATGCTCAGTGAAGATGTCAGTGTGAACCAAGAGAGAGAGAGAGAGAGAG 3748
Qy 961 GluSerLeuGlnGlyLysAspAspTyraGlyPheSerGlySerGlnLysLeuGluGlyGlu 980
Db 3749 GAGAGTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3808

QY 981 PheIysGlnProIleProAlaSerSerMetProGlnGlnIleGlyIleProThrSerSer 1000
DB 3809 TTCAAACAAACCAATTCCTGGCTTCACAGCCACAGAAATAGCAATCTTACAGATTC 3868
QY 1001 LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGluSer 1020
DB 3869 TTAACCTCAATGTTGATCTTCGGGGAAGCGGTTTATAGTAGTCTGTGGCCAAAGC 3928
QY 1021 ArgLeuArgGluSerIlyValPheProSerGluIleThrAspThrValAlaIleSerThr 1040
DB 3929 CGATTTCAGAAATCAAAAGTTTCCCAAGAAATACAGATACAGTTGCTGGCTCTACA 3988
QY 1041 AlaGlnSerProGlyMetLeuSerSerHisSerAlaSerSerLeuSerLeuGlnAla 1060
DB 3989 GCTCAGAGCCCTGSAATGAATGCTTCACCTGCATCATCTTACTTACACAGGCC 4048
QY 1061 PheSerGluLeuArgArgAlaGlnMetThrGluGlyProAsnThrAlaProProAsnPhe 1080
DB 4049 TTTTCTGAATCTTAGAGTGCCTCAATATGACAGAAAGACCCAAACAGCACTCCAACTTT 4108
QY 1081 SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal 1100
DB 4109 AGTCATACAGAGCAACCAATTTCCAGTAGTACCTCTTCTTAAGTACATTTGCTGAGTC 4168
QY 1101 ProThrThrAlaAlaIleThrAlaProValProAlaThrSerSerProProAsnAspIle 1120
DB 4169 CCACTACAGACAGACGACCAAGCCAGCTCCCTGCAACAAAGACCCCTCTTAATGACAT 4228
QY 1121 SerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGlyVal 1140
DB 4229 TCCACATCAGTAATTCAGTCTGAGGTACAGTGCCTCAAGAGGGAGATTGCTGAGT 4288
QY 1141 AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer 1160
DB 4289 GCCACCAAGCAAGGTGGTAACTTCAGGTGCTCCCATACCACTGGTGTGAAATCA 4348
QY 1161 ProValLeuSerSerValIleSerSerIleThrIleProAlaValIleSerIleSerThr 1180
DB 4349 CCAATATCTTCCAGCGTAGTTTCAAGTATACATATCTGCAAGTTGCTCAATATCTACT 4408
QY 1181 ThrSerProSerLeuGlnValProThrSerThrSerGluIleValIleSerSerThrAla 1200
DB 4409 ACATCCCGTCACATTCAGTCCCAATCCACATCTGAGATCGTTGTTCTAGTACGCA 4468
QY 1201 LeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlySerThrAlaThr 1220
DB 4469 CTGATATCTTCAGTAAAGTTCAGAACTTCAGCTCTGCAAGGGCAGATCTGATACC 4528
QY 1221 ProGlyProIlyProProAlaValIleSerGlnAlaIleAlaGlySerThrThrValGly 1240
DB 4529 CCAAGTCTTAAGCTCCAGCTGATGATCTCAGACGACGACGACGACACTACTGTGGGA 4568
QY 1241 AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer 1260
DB 4589 GCCACATTAACATCAGTTTCCACACCACTTCATTCACCAAGCAACAGCTTCACACTGTC 4648
QY 1261 IleGlnLeuSerSerSerThrThrProThrLeuAlaGluThrValValIleSerAla 1280
DB 4649 ATTAGGCTTAGCAGACAGTACTTACTCTCTTACTTACTGTAACCGTAACCGTGTAGGCA 4708
QY 1281 HisSerLeuAspIlyThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla 1300
DB 4709 CACTCACTAGTAAAGACATCTCAGACGATACAACTGATTTGGCTTCTCTCTGCA 4768
QY 1301 ProSerSerSerSerSerProGlyAlaGlyValSerSerIlyIleSerGlnProGlyGly 1320
DB 4769 CCACTTCCTCTCTCTCTCTCTGAGACGAGAGTGTCTAGTTATTTCTCAGCCCTGGGG 4828
QY 1321 LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla 1340
DB 4829 CTGATATCTTGGTCAATTCATGATGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4888
QY 1341 GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro 1360
DB 4889 GGAACCTACTTCTACACTTTATTATCCCAAGTACTAGTATCCCACTCTGTATGACGCT 4948
QY 1361 ValAlaAsnValProAlaValGlnIleThrLeuIleHisSerGlnProGlnProAlaLeu 1380
DB 4949 GTTCCAAATGCTGCTCTGTACAGCAGACACTAATTCATAGTACGCTCAACCAAGCTTGG 5008
QY 1381 LeuProAsnGlnProHisThrHisGlyProGluValAspSerAspThrGlnProIlyAla 1400
DB 5009 CTTCACCAAGCCCACTACTCATTTGCTGAACTGATTTCTATACAAACCAAGCT 5068
QY 1401 ProGlyIleAspAspIleThrThrLeuGluGluValLeuArgSerLeuPheSerGluHis 1420
DB 5069 CTTGGAATGATGACATTAAGAATTAAGAAAGAGAGTGGCTCTGTTCAAGTAAACAC 5128
QY 1421 SerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThrSerLeuValIleGluSer 1440
DB 5129 AGCTCATCTGGAGCTCAGCATGCTCTGTCTCCTGAGACCTCAGTACATAGAGAGC 5188
QY 1441 ThrValThrProGlyIleProThrThrAlaValAlaProSerIlySerLeuThrSerThr 1460
DB 5189 ACTGTCAACCAAGCAATCCCACTACTGCTGTGCAACCAAGCAACTCTGACTTCTACC 5248
QY 1461 ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr 1480
DB 5249 ACAAGTACTGTTTACCAACCAATTTACACTAGAAACAGTGTCTTGCAGATTACA 5308
QY 1481 ProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyValIly 1500
DB 5309 CCAAGTGTACACCTGGGCAAGTTTCTACCCAGTACAGCACTCATATAGAGAGTAAA 5368
QY 1501 ProGlyIlyThrAlaProSerIlyProProLeuThrIlyAlaProValLeuProValGlyThr 1520
DB 5369 CTTGGAATCTGCTCTCTCAACCACTCTTACTTACAGTCTCGGTGCTGCAAGTGGTACT 5428
QY 1521 GluLeuProAlaGlyThrLeuProSerGluGlnLeuProProPheProGlyProSerLeu 1540
DB 5429 GAACTTCACAGTACTCTTACCAAGCAAGTGTGCACTTCTTCCAGAGACCTTCTCTA 5488
QY 1541 ThrGlnSerGlnIleProLeuGluAspLeuAspAlaGlnLeuArgGlyThrLeuSerPro 1560
DB 5489 ACCAGTCCCAAGCACTCTTACAGAGATCTTGAAGTCTCAATTTGGAAGAAACCTTATG 5548
QY 1561 Glu***IleThrValThrSerAlaValGlyProValSerMetAlaIleProThrAlaIle 1580
DB 5549 GAATATATCAAGTACTTCTGCGTTGGTCTGTCTGTGCAATGGCGGCTTCAACGCAATC 5608
QY 1581 ThrGluAlaGlyThrGlnProGlnIlyValSerGlnValIlyGluGlyProValLeu 1600
DB 5609 ACRGAACGAGAAACACAGCTCAGAAAGGTGTTCTCAAGTCAAAAGAAAGGCCCTGTCTTA 5668
QY 1601 AlaThrSerSerGlyAlaGlyValPheIleMetGlyValArgPheGlnIleSerValAlaAla 1620
DB 5669 GCAACTATGTCAGAGCTGTGTGTTTAAATGAGTGAAGATTTCAAGGTTTCTGTTCAGCA 5728
QY 1621 AspGlyAlaGlnIlyGluGlyIlyAsnIlySerSerGluAspAlaIlySerValHisPheGlu 1640
DB 5729 GACGCTGCCCAAGAAAGAGGTAAATAATGATCAGAAATGCAAAAGTCTGTTCAATTTGAA 5788
QY 1641 SerSerThrSerGluSerSerValLeuSerSerSerSerProGluSerThrLeuValIly 1660
DB 5789 TCCAGCACTCAGAGCTCTCAGTCTATCAAGTGTGTGCTCAGAGAGTACTTGTGTAAA 5848
QY 1661 ProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGluSerAlaHis 1680
DB 5849 CCAAGGCCGAATGCAATACCATCTCGTATCTCTTCAAGTGTGCAAGAGTGGCCAC 5908
QY 1681 IlyThrThrAlaSerGluAlaIlySerSerAspThrGlyGlnProThrIlyValGlyArgPhe 1700
DB 5909 AAAACTACTGCTCAGAGGCAAGTCAACACTGGCAGCTTACCAAGGTTGAGCTTTT 5968
QY 1701 GlnValThrThrAlaAsnIlyValGlyValArgPheSerValSerIlyThrGluAspIly 1720

Db 5969 CAGGTGACAACTAAGCAAAAGTGGTCTTCTCTGATCAAAAACTGAGGACAG 6028
 Qy 1721 ILeThrApThrLysGlyGlyProValAlaSerProPheMetAspLeuGln 1740
 Db 6029 ATCACTGACACAAAGAAAGAGACAGTGCATCTCTCTTTATGGATTGGACAA 6088
 Qy 1741 AlaValLeuProAlaValIleProLysGlyGlyProGluLeuSerGluProSerHis 1760
 Db 6089 GCTGTTCTCTGCTGATGATCAAGAAAGAAAGAGAGCTGAACTGACAGACCTTCAT 6148
 Qy 1761 LeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspValAspAsnGly 1780
 Db 6149 CTAATATGGCCGCTTCTTGACCCGAGGCGCCCTTTTAACTAGAGATGATGATGT 6208
 Qy 1781 SerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsnLeu 1800
 Db 6209 TCCGGTAGTCCACACTGCCCATCAGTGAAGCTCAAGACCTTCTAGCCAAATCTA 6268
 Qy 1801 SerGlnSerLeuSerAsnSerPheAsnSerSerLysMetSerSerAspAsnGluSerAsp 1820
 Db 6269 AGTCAAGACCTTACTTAATTCATTAACTCTTCACTAGAGTACGCAATGAGTCAAGAT 6328
 Qy 1821 ILeGluAspGluAspLeuLysLeuGluLeuArgArgLeuArgAspLysHisLeuLysGlu 1840
 Db 6329 ATCGAAGATGAGAGCTTAAGATTAAGCTGACGACACTACGAAATAAACATCTCAAGAG 6388
 Qy 1841 ILeGlnAspLeuGlnSerArgGlnLysHisGlnIleGlnSerLeuLysYThrLysLeuGly 1860
 Db 6389 ATTACAGAGCTCGAGAGCTCGCCAGAGATGAATTAATCTTTGATACCAAACTGGGC 6448
 Qy 1861 LysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArgArgArgArg 1880
 Db 6449 AAGGTGCCCCCTGCTGATTAATATCCCCAGCTCTCTCTTCAAGGAGAAAGACGACGA 6508
 Qy 1881 ProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsnLysSerPro 1900
 Db 6509 CCCCTAAAGCAAAAGGAGAAATCTAGTGAAGAGCTCTTGAGGAAATAAAGCCCC 6568
 Qy 1901 GlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisProGlnGlnThr 1920
 Db 6569 CAGCTTTCAGGTAACTGCTGCTGCTCAGAGTGCACCTTCAAGCTTGCACCCCGACAGACC 6628
 Qy 1921 LeuHisProProGlyAsnIleProGluSerGlyGlnAsnGlnLeuLeuGlnProLeuLys 1940
 Db 6629 CTCACACCTCTCGCAACATCCAGAGCTCGGCGAGATCAGCTTAACGCCCTTAAG 6688
 Qy 1941 ProSerProSerSerAspAsnLeuLysSerAlaPheThrSerAspGlyAlaIleSerVal 1960
 Db 6689 CCATCTCCCTCCAGTGAACCTTATTCAGCTTCAACAGTGAATGGTGCATTTCAGTA 6748
 Qy 1961 ProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGlyAlaThrVal 1980
 Db 6749 CCAAGCTTCTGCTCCAGTCAAGGAAACGAGCAACCAACACCTGTGGGGCAACAGTG 6808
 Qy 1981 AsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLysGlyThrPhe 2000
 Db 6809 AACGCGCAAGCGCCCAAGCTCAGCTCTCTGCTGATGCTCCAGACGAAAGGACACATTC 6868
 Qy 2001 ThrAspAspLeuHisLysLeuValAspAsnThrAlaArgAspAlaMetAsnLeuSerGly 2020
 Db 6869 ACAATATACCTTGCAAGATGTGATGACAAATGGGCCGCAATGCAATCTTCCAGGC 6928
 Qy 2021 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 2040
 Db 6929 AGGAGAGAGCAAGGAGGACATGATTAAGAGGGCCCTGGAATGGCAAGAAATTTCTCT 6988
 Qy 2041 AlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGlyGlySerAlaProIleSer 2060
 Db 6989 GCACCTGGGCAACGTGATCTCCATGACCTCGAAACCTGGGTGCTGCCCCCATCTCT 7048
 Qy 2061 AlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProProGlnGlnThr 2080
 Db 7049 GCAAGCATAGCTACTCTCTAGTCACTTCAACAAAGTCTATGTGCCCCCAACAGCATAT 7108

Qy 2081 GlyPheProAlaThrProPheGlyAlaGlnThrSerGlyThrGlyGlyProAlaProGln 2100
 Db 7109 GGCTTTCAGCTACCCCATTTGGGCTGCTCAATGAGTGGACGGGTGGCCGACACACAG 7168
 Qy 2101 ProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsnIleSerAsn 2120
 Db 7169 CCACTTGGCCAGTTCACCTGTGGAACTGCGCTCTTGGAGAAATTTCAATCATCAGCAAT 7228
 Qy 2121 LeuGlnLysSerLysSerAsnProProGlySerAsnLeuArgThrThr 2136
 Db 7229 TTGCAGAAATTCATACGAAACCCCGAGCTCCACCTCGGAGCACT 7276
 RESULT 11
 US-10-767-471-454
 : Sequence 454, Application US/10767471
 : GENERAL INFORMATION:
 : APPLICANT: CARGILL, Michele et al.
 : TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 : TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
 : FILE REFERENCE: C0001505
 : CURRENT APPLICATION NUMBER: US/10/767,471
 : CURRENT FILING DATE: 2004-01-30
 : NUMBER OF SEQ ID NOS: 50231
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 454
 : LENGTH: 9639
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: (1)..(9639)
 : OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
 US-10-767-471-454
 Alignment Scores:
 Pred. No.: 0 Length: 9639
 Score: 10567.50 Matches: 2096
 Percent Similarity: 98.13% Conservative: 0
 Best Local Similarity: 98.13% Mismatches: 11
 Query Match: 97.74% Indels: 29
 Gaps: 1
 US-10-010-720-14 (1-2136) x US-10-767-471-454 (1-9639)
 Qy 1 MetSerGlyGlyAlaAlaGlnLysGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 956 ATGCTGCGCGCGCGCGAGAGAGACAGACGACCTCCGGTCTCTGCTCTCGCGC 1015
 Qy 21 ProAlaProAlaProLysAsnGlySerSerSerSerSerValGlyGlyLysLeuGly 40
 Db 1016 CCGGCTCTCTCGCCCAAGAAATGCTCCAGCTCCGATTCCTCCGTGGGGGAAACTGGGA 1075
 Qy 41 AlaAlaAlaAlaAspAlaValThrGlyArgThrGlnLysArgArgArgHisThr 60
 Db 1076 GCCGCGCGCGCGCGCTGTGACCGGCGAGACCGAGAGATCAAGCGCCCGCCGACACT 1135
 Qy 61 MetAspLysAspSerArgArgValAlaAlaAlaThrThrThrThrArgHisPhePhe 80
 Db 1136 ATGACAAAGACACCGTGGGGCGCGCGCACCACTACCACTAGAGACCGGCTCTTC 1195
 Qy 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 Db 1196 CGCGGAGCGCTCATCTGMACTCCAAAGCCACTGAGAGCTTCCCGGCTTCTCTT 1255
 Qy 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGlnProHis 120
 Db 1256 TCCCTGCCCAAGCCCAACATCCCGCGCTGTCCCGAGAGTGTCCACCGAGACCCAC 1315
 Qy 121 ArgGlnGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 1316 CGGAGAGAGACCGTGAAGCCGACCGCACTTCCAGGTAGCCGAGACCTCCAGCGCT 1375

QY 141 AlaAlaProGluGluGlnAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 1376 RCGCCCTGGGGAACAGCCGCTGCGGGCCCTCCCTCGACTGTCGCCAGAGTACC 1435
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProPro 180
 Db 1436 AGCAAAACCGCCAGTGTCCAGCCTTAGCTTGGGGAGCAAAAGAGACCCCGCG 1495
 QY 181 AlaArgSerGlySerGlyGlySerAlaLysGluProGlnGluLysSerGln 200
 Db 1496 GCGAGAACTGGCAGCGGCGCGAGCGCCAGAGCCAGAGAGAACGAGCTCAG 1555
 QY 201 GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
 Db 1556 CAGATGATATCGAAGAGCTGAGACCAAGCCCTGGGAATGTTAAAGATGCCGCTTT 1615
 QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
 Db 1616 CTCAAGTTTGACATCGAAATCGCAGAGGCTCTTTAAAGCGTCTCAAAAGTCTGAG 1675
 QY 241 ThrGluThrThrValGluValAlaTyrCysGluLeuGlnAspArgLysLeuThrLysSer 260
 Db 1676 ACTAAACACACCGTGAAGTGTCTGTGTGACTGACGAGATCAAAATTAACAAAGTCT 1735
 QY 261 GluArgGlnArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAsnIle 280
 Db 1736 GAGAGCAGAGATTTAAAGAGAGCTGAATGTTAAAGTCTTCAGCATCCCAATTT 1795
 QY 281 ValArgPheTyrAspSerTyrGluSerThrValLysGlyLysLysCysIleValLeuVal 300
 Db 1796 GTTGATTTTATGATTCCTCGGAAATCCACAGTAAAGAAAGATCATGTTTGTTG 1855
 QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrTyrLeuLysArgPheLysValMetLys 320
 Db 1856 ACTGACATTTATGACGCTGGAACCTTAACCGTATCTGAAGAGTTTAAAGTATGATGAG 1915
 QY 321 IleLysValLeuArgSerTyrCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
 Db 1916 ATCAAGTTTCAAGAGCTGGTGCCTGCAAGTCTTAAAGTCTTCAGTTTCTTCATCT 1975
 QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 Db 1976 CGAATCCACCTTATCATTCACCGGATCTTAATGTACAAACATCTTATCACCGGCT 2035
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuLeuAlaThrLeuLysArgAlaSerPheAla 380
 Db 2036 ACTGGCTGAGTCAAGATTTGAGACCTCGTGTGCAACCTGAAGCGGCTCTTTGGCC 2095
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetTyrGluGluLysTyrAsp 400
 Db 2096 AAGAGTGTGATAGTACCCCAAGTTCATGGCCCTGAGATGTATGAGGAATATGAT 2155
 QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
 Db 2156 GAATCCCTTACGTTATGCTTTGGATGTGATGCTTGAGATGGCTTACATCGATATAT 2215
 QY 421 ProTyrSerGluCysGlnAsnAlaIleGlnIleTyrArgArgValThrSerGlyValLys 440
 Db 2216 CTTTACTCGAGTGCCTCAAAATGTGCGACATTTACCGTGTGCGACAGGGGGTAAAG 2275
 QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
 Db 2276 CCAAGCATTTTGAACAAGTACCAATTCCTGAATGAGGAATATTTGAAGATGATATA 2335
 QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
 Db 2336 CGCAAAACAAAGATGAAGATATTCATCAAAAGCTTTGAACCAATGCTTTTCCAA 2395
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluGluAspAspGlyGluLysIleAlaIle 500
 Db 2396 GAGGAAACAGAGAGTACGGGTGAGATTTAGCAAGAAAGATGAGAAAAAATAGCCATA 2455
 QY 501 LysLeuTyrPheLysArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520

Db 2456 AAATTAAGCTACTACTATTGAAGATTTTAAGAAATTTAAAGGAAAAATCAAGATATGCA 2515
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 Db 2516 GCTATTGAGTTTCTTTTGAATTTAGAGAGATGTCCAGAAAGATGTGCAACAAGAAATG 2575
 QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 Db 2576 GTAGAGTCTGAGTATGTCTGTGAAGGTGATCAAGACCATGGCTTAAAGCTATCAAGAC 2635
 QY 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGluGlu 580
 Db 2636 AGAGTATCATTTAATTAAGAGAAACGAGACAGGCGAGTTGTCACGAGAGACACAGA 2695
 QY 581 LysLysLysGluGluGluSerSerLeuLysGlnValGluGlnSerAlaSerGln 600
 Db 2696 AAAAAAAGCAGAGAGACAGTCTCAACAGCAGGTAGAACATCCAGTCTTCCAG 2755
 QY 601 ThrGlyIleLysGluLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 Db 2756 ACAGAAATCAAGCAGCTCCCTTGTGTAGCACCGCATACCTACTGCTTACCACTTCA 2815
 QY 621 AlaSerValSerThrGlnValGluProGluGluLysAspGlnHisGlnGlnLeu 640
 Db 2816 GCTTCACTTCTTACACAGATGAACTGAAAGAACTGAGCAGATCAACATCAACACTA 2875
 QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 Db 2876 CAGTACCAAGCAACCCAGTATATCTGTGTATCTGATGGACGGTTGACAGTGTCAAGGA 2935
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 Db 2936 TCCTCTCTTTCAYAGATCTCGATGAGAGCCCAAGCAACAGTTCATATGATGCCAA 2995
 QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 Db 2996 CATGAACAGCACTTTTACAGGACAGTCCAGGGCATATACCTTCTACTGTCCAAAGCA 3055
 QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 Db 3056 CAGTCTCAGCCCATGAGGTATATCCACCTCAAGTGTG----- 3094
 QY 721 GlnProSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
 Db 3094 ----- 3094
 QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
 Db 3095 -----CAGGGAATPACGACAGACAGCCCTCCCTCAACAGACAGTGCAGTATTCATTTCA 3148
 QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 Db 3149 CAGACATCAACCTCCAGTGGGCGACACTACGACAGCAGCTGATGATCAACCTCAAGCTCCA 3208
 QY 781 GlnValLeuProGlnValSerAlaGlyLysGlnSerThrGlnGlyValSerGlnValAla 800
 Db 3209 CAACTCTTGCTCAAGTATACGCTGGAAACAGATCTCAAGGAGTCTTCAGGTGCT 3268
 QY 801 ProAlaGluProValAlaValAlaGlnProGlnAlaThrGlnProThrThrLeuAlaSer 820
 Db 3269 CTGCAAGAGCAGATTTGAGTAGCAGACAGMCCCAAGCTACCCAGCGCACTTTGGCTTCC 3328
 QY 821 SerValAspSerAlaHisSerAspValAlaSerGlyMetSerAspGlyAsnGluAsnVal 840
 Db 3329 TCTGTAGACGTGCACATTCAGATGTGCTTCAGGTATGATGCAATGAGAACGTC 3388
 QY 841 ProSerSerSerGlyArgHisGluGlyArgThrThrLysArgHisTyrArgLysSerVal 860
 Db 3389 CCAITCTTCACTGGAGAGCATGAAAGAGAACTTCAAAAGGCACTTACCAAAATCTGTA 3448
 QY 861 ArgSerArgSerArgHisGlyLysThrSerArgProLysLeuArgIleLeuAsnValSer 880

Db	3449	AGGAGTCCTCTGACATGAAAAAATTACGCCCAAAATTAAGATTGTTTCA	3508
Qy	881	AsnLysGlyAspArgValValIguCyGlnLeuGluThrHisAsnArgLysMetValThr	900
Db	3509	AAATTAAGAGACCGAGTAGTAGAATGTCAATTAGAGCTCATATATAGAAAAATGTTTACA	3568
Qy	901	PhelLysPheAspLeuAspGlyAspAsnProGluIuLeaIaThrIleMetValAsnAsn	920
Db	3569	TTCAAAATTTGACCTAGTAGGTGACAAACCCGAGAGAGATACCAAAATTAATGAGTGAACAT	3628
Qy	921	AspPheIleLeuAlaIleGluArgGluSerPheValAspGlnValArgGluIleIleGlu	940
Db	3629	GACTTATTTTACGATATAGAGAGAGAGTGTGTTGTGATCAAGTGCAGAAATTAATTTGAA	3688
Qy	941	LysAlaAspGluMetLeuSerGluAspValSerValGluProGluGlyAspGlnGlyLeu	960
Db	3689	AAAGCTGATGAAATGTCTCAGTAGAGAGTGTGAGTGGAAACCAAGGTTGATCAGGATTTG	3748
Qy	961	GluSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySerGlnLysLeuGluGlyGlu	980
Db	3749	GAGAGTCTACAAAGAAAGATGACTATGCTTTTCAGGTTCTCAGAAATTGAAAGAGAG	3808
Qy	981	PheLysGlnProIleProAlaSerSerMetProGluGlnIleGlyIleProThrSerSer	1000
Db	3809	TTTAAACAAACCAATTCCTGCTCTTCATGCAACAGCAAAATAGGCATTCCTACCATTTCT	3868
Qy	1001	LeuThrGlnValValHisSerAlaGlyArgArgPheIleValSerProValProGluSer	1020
Db	3869	TTAACTCAAGTTTTCATTTCTGCGGAGAGCGGTTTATAGTAGTCTGTGCGCAAGAAAGC	3928
Qy	1021	ArgLeuArgGluSerLysValPheProSerGluIleThrAspThrValAlaAlaSerThr	1040
Db	3929	CGATTTCGAAATCAAAAGTTTCCCGAGTAATACATACATACAGTGTGCTGCTCTACAC	3988
Qy	1041	AlaGlnSerProGlyMetLeuSerHisSerHisSerAlaSerSerLeuGlnGlnAla	1060
Db	3989	GCTCAAGCCCTGAAATGAACTGTCTCACTGTGCATCATCTTATGTTCAACACAGGCC	4048
Qy	1061	PheSerGluLeuValArgArgAlaMetThrGluGlyProAsnThrAlaProProAsnPha	1080
Db	4049	TTTTTGAAGTTTAGAGTGCCTCAATGACGAAGAGACCAAAACAGCACCCTCCAACTTT	4108
Qy	1081	SerHisThrGlyProThrPheProValValProProPheLeuSerSerIleAlaGlyVal	1100
Db	4109	AGTCATACAGAGCAACATTTCCAGTAGTACTCTTTCTTAAGTAGCATTTGCTGAGATC	4168
Qy	1101	ProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAlaIle	1120
Db	4169	CCAACTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	4228
Qy	1121	SerThrSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGlyVal	1140
Db	4229	TCCACTACAGTATTCAGTCTGAGGTTTACGTGCCCATGAAAGGGAGATGCTGAGATT	4288
Qy	1141	AlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGluSer	1160
Db	4289	GCCACACAGACAGAGTGTGTAACCTCAGAGTGTCTCCCATACCACTGTGTGATCA	4348
Qy	1161	ProValLeuSerSerValValSerSerIleThrIleProAlaValValSerIleSerThr	1180
Db	4349	CAAGTACTTTCACGGCTAGTTTCAAGATATCAAAATCTGCACTTCTCAAAATATCTACT	4408
Qy	1181	ThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThrAla	1200
Db	4409	AACATCCCGTCACTTCAAGTCCCAATCCACATCTGAGATGCTTGTTCATATACAGCA	4468
Qy	1201	LeuTyrProSerValThrValSerAlaThrSerAlaSerAlaGlyLysSerThrAlaThr	1220
Db	4469	CTGTATCTCTTCAGTATCAGTTTCAGCACTTCAGCTTCAGGGGGAGTACTGCTACCC	4528
Qy	1221	ProGlyProLysProProAlaValValSerGlnGlnAlaAlaGlySerThrThrValGly	1240
Db	4529	CCAGGTCTTAAGCTTCAAGCTGTAGTATCTCAGCAGGACGACAGGCACTACTGTGGCA	4588

Qy	1241	AlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeuSer	1260
Db	4589	GCCACTTAATACATCACTTTTCTACACACACTTCAATTCACAGCAGAGCTTACAGCTTSC	4648
Qy	1261	IleGlnLeuSerSerSerThrSerThrProThrLeuAlaGluThrValValIleSerAla	1280
Db	4649	ATTCAGCTTAGCAGCAGTACTTCTACTCTCACTTAGCTGAACCGGTGATGAGCGCA	4708
Qy	1281	HisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSerAla	1300
Db	4709	CACCTCACTAGATTAAGACATCTCATACAGTACAACTGAGATGGCTTCTCTCTCTGCA	4768
Qy	1301	ProSerSerSerSerSerProGlyAlaGlyValSerSerTyrIleSerGlnProGlyGly	1320
Db	4769	CCATCTTCTCTTCTCTCTCTCTGAGAGAGAGTCTTAGTTATATTTCTCAGCTGTGGG	4828
Qy	1321	LeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAlaAla	1340
Db	4829	CTGCATCTTTGGTCAATTCATCAGATGATGATTTACTTCTTCTTCTTCTTCTTCTTCT	4888
Qy	1341	GlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGlnPro	1360
Db	4889	GGACCTACTTCTACACCTTTATTTACCCCAAGTACCTAGATCCACCTGTGACAGCCT	4948
Qy	1361	ValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAlaLeu	1380
Db	4949	GTTGCCAATGTGCTGCTCTGACGACGACATCATTTATGATAGCTTCAACAGCTTTG	5008
Qy	1381	LeuProAsnGlnProHisThrHisCysProGluValAspSerAspThrGlnProLysAla	1400
Db	5009	CTTCCCAACAGGCCCATCATCTATGCTCTGAAATGATTTCTGATACAAACCAAGCT	5068
Qy	1401	ProGlyIleAspAspIleLysThrThrLeuGluGluLysLeuArgSerLeuPheSerGluHis	1420
Db	5069	CTGGAATGATGACATTAAGACTTAGAAGAAACCTGCGTCTGTTCAGTGAACAC	5128
Qy	1421	SerSerSerGlyValAlaHisAlaSerValSerLeuGluThrSerLeuValIleGluSer	1440
Db	5129	AGCTACTGTGAGCTCAGATGATGCTGTCTCTCACTGAGAACCTCACTAGTACAGAGGC	5188
Qy	1441	ThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSerThr	1460
Db	5189	ACTGTACACAGGACATCCCACTACTGCTGTGACCAAGCAAACTCTGACTTCTAC	5248
Qy	1461	ThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProValThr	1480
Db	5249	ACAACTACTTGTCTTACACCAACCAATTTTACCACTAGGAACAGTTGCTTGCACAGTTTACA	5308
Qy	1481	ProValValThrProGlyGlnValSerThrProValSerThrThrSerGlyValLys	1500
Db	5309	CCAGTGTACACCTGGGCAAGTTTCTACCCAGTAGACATCTACATCAGAGAGTGA	5368
Qy	1501	ProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGlyThr	1520
Db	5369	CTGTGAATCGTCCCTCCCAAGCACCTTAACTAAGGCTCCGGTGTGCAAGTGGTACT	5428
Qy	1521	GluLeuProAlaGlyThrLeuProSerGluGlnLeuProProPheProGlyProSerLeu	1540
Db	5429	GAACTTCCAGCAGTACTTACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	5488
Qy	1541	ThrGlnSerGlnIleProLeuGluAspLeuAspAlaGlnLeuArgThrThrLeuSerPro	1560
Db	5489	ACCCAGTCCACAGCACTTACAGATCTTGAATGCTCAATTTAGAGAAACATTAAGTCCA	5548
Qy	1561	Glu***IleThrValThrSerAlaValGlyProValSerMetAlaAlaProThrAlaIle	1580
Db	5549	GAGATATATCAGATGATCTTGTGGGTGTCTGCTGATGATGAGGAGCTCAACAGCAATC	5608
Qy	1581	ThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGlnGlyProValLeu	1600
Db	5609	ACRGAGCAGGAAACAGGCTCAGAAAGGAGTGTCTTCAAGTCAAAAGAGGCTGTGCTTA	5668

Db 1016 CCGGCTCTGCCCCCAAGATGCTCCAGCTCCGATTCCTCCGCGGAGGAAACTGGGA 1075
 QY 41 AAlaAlaAlaAlaAlaAlaValThrGlyArgThrGluTyrArgArgArgGlnSerThr 60
 Db 1076 GCGCGGCGCGCGCGCGCTGTGACCGGAGAGGACGAGAGATACAGCGCGCGCGCGCAACT 1135
 QY 61 MetAspLysAspSerArgGlyAlaAlaAlaThrThrThrThrThrGluHisArgPhePhe 80
 Db 1136 ATGACAAAGACAGCCGTGGGGCGGCGGACCACTACCACTGAGCACCGCTTCCTTC 1195
 QY 81 ArgArgSerValIleCysAspSerAsnAlaThrAlaLeuGluLeuProGlyLeuProLeu 100
 Db 1196 CGCGGAGGCTCATCTGAGACTCCATGCACTGCACTGAGAGCTTCCGGGCTTCCTCT 1255
 QY 101 SerLeuProGlnProSerIleProAlaAlaValProGlnSerAlaProProGluProHis 120
 Db 1256 TCCCTGCGCCAGCCAGCATCCCGGGGCTGCCGAGAGTGCTCCACCGGAKCCCGAC 1315
 QY 121 ArgGluGluThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 1316 CCGGAAAGACCGGTGACCGGACCGGCACTCCAGGTAGCCAGAGCTCCAGCGGCT 1375
 QY 141 AlaAlaProGlyGluGluAlaValAlaGlyProAlaProSerThrValProSerSerThr 160
 Db 1376 RCCGCCCCCTGGGAGACAGGCGCTCGCGGCGCTGCCCCCTCGACTGTCCAGAGTACC 1435
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGluGluProProPro 180
 Db 1436 AGCAAAAGACCGCCAGGTGCCAGCTTAGCTTGCGGAGCAAAAGAGGAGCGCGCGCG 1495
 QY 181 AlaArgSerGlySerGlyGlyGlySerAlaLysGluProGlnGluLysArgSerGlnGln 200
 Db 1496 GCGAGAGGTGCGAGCGCGCGGCGGACGCGCAAGAGAGCAAGAGAAACCGAGCCAGCG 1555
 QY 201 GlnAspAspIleGluGluLeuGluThrLysAlaValGlyMetSerAsnAspGlyArgPhe 220
 Db 1556 CAGGATGATTCGAAGAGCTGAGACCAAGGCGGTGGAAATGCTAAACGATGCGCTT 1615
 QY 221 LeuLysPheAspIleGluIleGlyArgGlySerPheLysThrValTyrLysGlyLeuAsp 240
 Db 1616 CTCAGATTGACATCGAAATCGGACAGAGCTCCTTTAAGCGGTCTCAAAAGCTCGGAC 1675
 QY 241 ThrGluThrThrValGluValAlaATPCysGluLeuGlnAspArgLysLeuThrLysSer 260
 Db 1676 ACTGAACCAACCGTGGAAGTGGCTGTGTGAACCTGAGGATCGAAATTAACAAGTCT 1735
 QY 261 GlnArgGluArgPheLysGluGluAlaGluMetLeuLysGlyLeuGlnHisProAlaIle 280
 Db 1736 GAGAGCGACAGATTTAAAGAGAAAGCTGAATGTAAAGGTCTTCAAGCATCCCAATATT 1795
 QY 281 ValArgPheLysAspSerTrpGluSerThrValLysGlyLysLysCysIleValLeuVal 300
 Db 1796 GTTAGATTATATGATTCCTGGGAATCCACAGTAAAGAAAGAAAGTGCATGTTTGGTG 1855
 QY 301 ThrGluLeuMetThrSerGlyThrLeuLysThrThrLeuLysArgPheLysValMetLys 320
 Db 1856 ACTGAACCTATGACGTCTGGAACACTTAAACGATCTGAAAGGTTTAAAGTATGAAG 1915
 QY 321 IleLysValLeuArgSerTrpCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340
 Db 1916 ATCAAAAGTTCTAAGAGCTGGTGCGGTCAAGATCTTAAAGGTCTTCAAGTTCTTCAACT 1975
 QY 341 ArgThrProProIleIleHisArgAspLeuLysCysAspAsnIlePheIleThrGlyPro 360
 Db 1976 CGAAGCTCCACCTATCATTCACCGCGATCTTAAATGTGACAACTTTATCACCGGCGCT 2035
 QY 361 ThrGlySerValLysIleGlyAspLeuGlyLeuAlaThrLeuLysArgLysAspPheAla 380
 Db 2036 ACTGGCTCAGTCAAGATGTGAAGACTCGGTCTGGCAACCTGAAAGCGGCTCTTTGGC 2095
 QY 381 LysSerValIleGlyThrProGluPheMetAlaProGluMetThrTrpGluLysTyrAsp 400
 Db 2096 AAGAGTGTATAGTACCCCAAGATTCATGGCCCTGTAAGTGTATGAGAGAAATATGAT 2155

QY 401 GluSerValAspValTyrAlaPheGlyMetCysMetLeuGluMetAlaThrSerGluTyr 420
 Db 2156 GAATCCGTTACGCTTATATGCTTTGGGATGTGATGCTTAGAGTGCATCATCTGAATAT 2215
 QY 421 ProTyrSerGluCysGlnAsnAlaAlaGlnIleTyrArgArgValThrSerGlyValLys 440
 Db 2216 CTTTACTCGGAGTGCAGAAATGCTGCRCAATCATCGTGGCTGACCAAGTGGGGTGAAG 2275
 QY 441 ProAlaSerPheAspLysValAlaIleProGluValLysGluIleIleGluGlyCysIle 460
 Db 2276 CCAGCAAGTTTGAACAAGTAGCAATTCGAAGTGAAGAAATTAATGAAGATGATATA 2335
 QY 461 ArgGlnAsnLysAspGluArgTyrSerIleLysAspLeuLeuAsnHisAlaPhePheGln 480
 Db 2336 CGACAAACAAATATGAATATTTCCATCAAAAGCCTTTGAACCATGCTTTTCCAA 2395
 QY 481 GluGluThrGlyValArgValGluLeuAlaGluLysAspGlyGluLysIleAlaIle 500
 Db 2396 GAGGAAACAGAGAGTACGGGTAGAAATTAGCAGAGAAAGAYATGAGAAATAATAGCATTA 2455
 QY 501 LysLeuThrLeuArgIleGluAspIleLysLysLeuLysGlyLysTyrLysAspAsnGlu 520
 Db 2456 AATTTATGGCTACGTATGAAATTAAGAAATTAAGGAAATAACAAAGATTAATGAA 2515
 QY 521 AlaIleGluPheSerPheAspLeuGluArgAspValProGluAspValAlaGlnGluMet 540
 Db 2516 GCTATTGAGTTTCTTTATGATTTAGAGAGAGATGCTCCAAATATGTGCAAGAAATG 2575
 QY 541 ValGluSerGlyTyrValCysGluGlyAspHisLysThrMetAlaLysAlaIleLysAsp 560
 Db 2576 GTAGAGTCTGGGTATGTCTGTGAAGGTGATCAACAAGCATAGCTTAAGCTATCAAGAC 2635
 QY 561 ArgValSerLeuIleLysArgLysArgGluGlnArgGlnLeuValArgGluGluGlnGlu 580
 Db 2636 AAGATATCATTAATTAAAGAGAAACGAGAGCAGCGAGTGGTGAAGGAGGAGCAAGAA 2695
 QY 581 LysLysLysGluGluGluSerSerLeuLysGlnGlnValGluGlnSerSerAlaSerGln 600
 Db 2696 AAAAAAACAAGAAAGACAGCACTCAAAACAGCAGTGAACAAATCCAGTCTTCCAG 2755
 QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
 Db 2756 ACAGGAATCAAGAGAGCTCCCTCTGCTAGCACCGGAGTACCTCATCTTACCACTTCA 2815
 QY 621 AlaSerValSerThrGlnValGluProGluGluProGluAlaAspGlnHisGlnGlnLeu 640
 Db 2816 GCTTCAGTTTCTTACACAAAGTAGAACTGAAGAACTGAGGCAAGATCAACATCAACAATA 2875
 QY 641 GlnTyrGlnGlnProSerIleSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
 Db 2876 CAGTACACCAACCCAGTATATCTGTTATCTGATGGAAGGTTTACAGTGTATCGAGGA 2935
 QY 661 SerSerValPheThrGluSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
 Db 2936 TCCCTGTCTTCAVAGAATCTGAGTGACAGCAACACACAGATTCATATGTTCCCA 2995
 QY 681 HisGluGlnAlaHisSerThrGlyThrValProGlyHisIleProSerThrValGlnAla 700
 Db 2996 CATGAACAGCACATTTCAAGGACAGACAGTCCAGGGCAATATCTTCTGCTCCAGGA 3055
 QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlyGlnSerGlnGly 720
 Db 3056 CAGTCTACGCCCAATGGGTTATATCAACCTCAAGTGTGGACAGGGGAGAGCCAGGAT 3115
 QY 721 GlnProSerSerSerSerLeuThrGlyValSerSerSerGlnProIleGlnHisProGln 740
 Db 3116 CAGCATCTCTCAAGTAGCTTAACAGGGGTTTATCTTCCCAACCATCAACATCTCTCAG 3175
 QY 741 GlnGlnGlnGlyIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
 Db 3176 CAGCAGCAGGAGATACAGCAGACAGCCCTCTCTCAACAGACAGTGCATTAATCACTTCA 3235

QY 761 GlnThrSerThrSerSerGluAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
 DB |||||
 QY 3236 CAGACATCAACCTCCAGTAGAGGCCACTACTGACACAGCCAGTAGTCACCTCAAGCTCCA 3295
 DB |||||
 QY 781 GlnValLeuProGlnValSerAlaGlyLeuGln----- 791
 DB |||||
 QY 3296 CAAGCTTGGCTCAAGATATCAAGCTGGAAAAACAGCTTCCAGTTCCAGGACGATACCAACT 3355
 DB |||||
 QY 791 ----- 791
 DB 3356 ATCCAAAGCGAACCTCAGATCCAGTTGCGACACAACCTCGGTTGTTCCAGTCCACTCT 3415
 QY 791 ----- 791
 DB 3416 GGTGCTCATTTCTTCCAGTGGAGACAGCCGCTCCCTAATCTCTTCTCTCCCTCAGTACCT 3475
 QY 791 ----- 791
 DB 3476 GTCTCATATTCCTCATATCACTCCATGTGTACAGGCTCAGACAGGTTTCTCATCC 3535
 QY 791 ----- 791
 DB 3536 CTTCCTCATCAATGCGACGTGCAGTATCTACGCTCTGCTCAGTTGGCTTCATCTGCT 3595
 QY 791 ----- 791
 DB 3596 ACAACAGCTGCGATCCCGGGGGATATCACTGTGTTCTTAGTCAAGCTTCCACCTTTCTG 3655
 QY 791 ----- 791
 DB 3656 CAGCTGTGACTCAGGCTGCCAAGTCAAGTTCCACCAAGCTCTCAACACGAGATTGAC 3715
 QY 791 ----- 791
 DB 3716 TCCATGGGAATACAGCTAATCTTGGACAAGCTGTGAGTTCCACTTCTCTGAGAT 3775
 QY 791 ----- 791
 DB 3776 GTTCTGTACAGGGCTTCCCACTCGACTGCCACACAGTACCCAGAGATTCAAATATT 3835
 QY 791 ----- 791
 DB 3836 GCTCCCTTTCCAAAGTGCTTCTGTGTGATCATCTTCAAGTCCATCCCTCCCATG 3895
 QY 791 ----- 791
 DB 3896 CCGACAGAGTACTGCTACAGCTGAGTACTTCCACAGTGTGAGAGCTTATGTGAA 3955
 QY 791 ----- 791
 DB 3956 TCAATCTTTTACTTCTTATGGTGTGTAGAGAGCAGTTCAAGTCTCCACGACGA 4015
 QY 792 ----- 792
 DB 4016 GGGAGTTTACACAAGCCCCCATACATCTCCAGACAGAGTTTGGAGATACAG 4075
 QY 795 GlyValSerGlnValAlaProAlaGlnProValAlaValAlaGlnProGlnAlaThrGln 814
 DB |||||
 QY 4076 GGAAGTCTCTCAAGGTGCTCTGACAGGCCAGTGTGACAGTACACAGMCCCAAGCTACCCAG 4135
 DB |||||
 QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
 DB |||||
 QY 4136 CCGAACCACTTTGGCTTCTCTGAGACAGTGCACATTCAGATGTGCTTCAAGTATGAGT 4195
 DB |||||
 QY 835 AspGlyAsnGluAsnValProSerSerSerGlyArgHisGluGlyArgThrThrIleArg 854
 DB |||||
 QY 4196 GATGGCAATAGAACGCTCCATCTTCCAGTGAAGCATGAGAGAACTACAAAAACGG 4255
 DB |||||
 QY 855 HisTyrArgIleSerValAlaArgSerArgSerArgHisGluGlySerThrSerArgProIleLeu 874
 DB |||||
 QY 4256 CATTACCGAAATCTGTAAAGAGTGGCTCTCGACATGAAAAAACTTCAACCCCAAAATTA 4315
 DB |||||
 QY 875 ArgIleLeuAsnValSerAsnLeuGlyAspArgValValGluGlySerGlnLeuGluThrHis 894
 DB |||||

DB 4316 AGAATTTTAAATGTTTCAATTAAGAGACCGAGTGTGATGATGTCAATTAAGACTCAT 4375
 QY 895 AsnArgIleMetValThrPheIysPheAspLeuAspGlyAspAsnProGlnIleAla 914
 DB |||||
 QY 4376 AATAGAAATGGTTCATTTCAATTAATTTGACTAGATGTGACACACCCAGAGAGATAGCA 4435
 DB |||||
 QY 915 ThrIleMetValAspAsnSerPheIleLeuAlaIleGluArgGluSerPheValAspGln 934
 DB |||||
 QY 4436 ACAATTTATGTAACATATCACTTATTTAGCAATAGAGAGAGATGCTTTGTGGATCA 4495
 DB |||||
 QY 935 ValArgGluIleIleGluIysAlaAspGluMetLeuSerGluAspValSerValGluPro 954
 DB |||||
 QY 4496 GTCCGAAATTTATTAAGAAAAAGCTGATGAATCTCAGTGAAGATTCAGTGGAAACA 4555
 DB |||||
 QY 955 GluGlyAspGlnGlyLeuGlnSerLeuGlnGlyLysAspAspTyrGlyPheSerGlySer 974
 DB |||||
 QY 4556 GAGGGTGATCAGGGATTTGAGAGATCTACAAAGAAAGATGACTATGGCTTTTCAAGGTTCT 4615
 DB |||||
 QY 975 GlnIleLeuGlnGluGluPheIysGlnProIleProAlaSerSerMetProGlnIle 994
 DB |||||
 QY 4616 CAGAAATTTGAAGAGAGATTCAAACCAATTCCTGCGCTTCCATGCGACAGCAAAATA 4675
 DB |||||
 QY 995 GlyIleProThrSerSerLeuThrGlnValValHisSerAlaGlyArgArgPheIleVal 1014
 DB |||||
 QY 4676 GGCATTTCTACCAAGTCTTTAATCAAGTTGTTCACTTCCGGGAGAGCGGTTATATG 4735
 DB |||||
 QY 1015 SerProValProGluSerArgLeuArgGluSerLysValPheProSerGluIleThrAsp 1034
 DB |||||
 QY 4736 AGTCTGTGCCGAAAGCCGATTAACAGATCAAAAGTTTCCCGAGTAAATAAACAGAT 4795
 DB |||||
 QY 1035 ThrValAlaAlaSerThrAlaGlnSerProGluIleMetAsnLeuSerHisSerAlaSer 1054
 DB |||||
 QY 4796 ACAGTTGCTGCTTCAAGCTCAGAGCCCTGGATATACCTGTCTCATCTGTGATATCC 4855
 DB |||||
 QY 1055 LeuSerLeuGlnGlnAlaPheSerGluLeuArgArgAlaGluMetThrGlnGlyProAsn 1074
 DB |||||
 QY 4856 CTTAGTCTCAACAGAGCCCTTTCTGTAACCTTAGACGTGCCAAATGACAGAGACCAAN 4915
 DB |||||
 QY 1075 ThrAlaProProAsnPheSerHisThrGlyProThrPheProValAlaProPheLeu 1094
 DB |||||
 QY 4916 ACAGCACTTCCAAACTTGTATGATACAGAGCAACATTTCCAGTATACCTCTTTCTTA 4975
 DB |||||
 QY 1095 SerSerIleAlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSer 1114
 DB |||||
 QY 4976 AGTAGATTGCTGAGATCCCAACACAGACGACGACACAGCAAGCTCTGACAGCAAGC 5035
 DB |||||
 QY 1115 SerProProAsnAspIleSerThrSerValIleGlnSerGluValThrValProThrGlu 1134
 DB |||||
 QY 5036 AGCCCTCTATATGACATTTCCATCATCAGTATTCAGTCTGAGGTTACAGTGCCCATGAA 5095
 DB |||||
 QY 1135 GlnGlyIleAlaGlyValAlaThrSerThrGlyValValThrSerGlyLeuProIle 1154
 DB |||||
 QY 5096 GAGGGATTTGCTGGAGTTGCCACACAGCAGAGTGTGTAATCTCAGTGTGCTCCCATAT 5155
 DB |||||
 QY 1155 ProProValSerGlnSerProValLeuSerSerValValSerSerIleThrIleProAla 1174
 DB |||||
 QY 5156 CAACCTGTGTGAAATCACACAGTACTTCCAGCGTGTGTTCAAGTATTCACAAATACCTGCA 5215
 DB |||||
 QY 1175 ValValSerIleSerThrThrSerProSerLeuGlnValProThrSerThrSerGluIle 1194
 DB |||||
 QY 5216 GTTGTCTCAATATCTATCATATCCCGTCACTTCAAGTCCCAATCCATCATCTGAGATC 5275
 DB |||||
 QY 1195 ValValSerSerThrAlaLeuTyrProSerValThrValSerAlaThrSerAlaSerAla 1214
 DB |||||
 QY 5276 GTTGTCTTCAAGTACAGACAGTATCTCTTCACTAGTAAAGTTCACAACTTCAAGCTCTGCA 5335
 DB |||||
 QY 1215 GlyGlySerThrAlaThrProGlyProIysProProAlaValValSerGlnGlnAlaAla 1234
 DB |||||
 QY 5336 GGGGGGAGTACTCTACCCAGGTTCTTAACTTCCAGCTCTAATATCTCAGCGGAGGCA 5395
 DB |||||
 QY 1235 GlySerThrThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSer 1254
 DB |||||

Dd	5396	GCGAGACATCACTGTGGGAGACCACATTACATCATGTTTCCACCAACTTCATTCCCAAGC	5455
Oy	1255	ThrAlaSerGlnLeuSerIleGlnLeuSerSerThrSerThrProThrLeuAlaGlu	1274
Dd	5456	ACAGCTTCACAGCTGTGSCATTGACGCTTAGCACAGATCTTCACTCCTACTTAGCTGA	5515
Oy	1275	ThrValValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeu	1294
Dd	5516	ACCGGGTAGTAGGCACAACTCACTAGTAAGAATCATCATGACAGTACAACACTGATTGG	5575
Oy	1295	AlaPheSerLeuSerAlaProSerSerSerSerProGlyAlaGlyValSerSerLys	1314
Dd	5576	GCTTTCTCCCTCTCTGACACATCTTCCTCTCTCTCTCCGAGGAGAGTGTCAGTAT	5635
Oy	1315	IleSerGlnProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrPro	1334
Dd	5636	ATTCTTCAGCCCGTGGGGGTGATCTTTGGTCATTCATGATAGTATAGCTTCACTCT	5695
Oy	1335	IleLeuProGlnAlaAlaGlyProThrserThrProLeuLeuProGlnValProSerIle	1354
Dd	5696	ATTCTTCCCCAAGCAGCAGGACCTTCACTTCACACTTATTATACCAGTACCTAGATAC	5755
Oy	1355	ProProLeuValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSer	1374
Dd	5756	CCACCTTGATACACCTGTGGCCATATGTGCTGTGACGACGACACTATATCATAGT	5815
Oy	1375	GlnProGlnProAlaLeuLeuProAsnGlnProHisThrHisCysProGluValAspSer	1394
Dd	5816	CAGCCTCAACACAGCTTGCTGCCAACACAGCCCACATCTATTGTCTCGAAGTATGATTC	5875
Oy	1395	AspThrGlnProLysAlaProGlyLysAspAspIleLysThrLeuGlnLysLeuArg	1414
Dd	5876	GATACACAAACCCAAAGCTCTGGATTTGATGACATAAAGACTCTTAAAGAAACACTGGGG	5935
Oy	1415	SerLeuPheSerGlnHisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGlnThr	1434
Dd	5936	TCTCTGTTCAGTAGACACAGCTCATCTGAGCTGACATGGCTCTGTCTCACTGGAAC	5995
Oy	1435	SerLeuValIleGlnLeuSerThrValThrProGlyLysLeuProThrAlaValAlaProSer	1454
Dd	5996	TCACTAGTCATAGAGAGACACTGTCCACACAGGACATCCCACTACTGCTGTGGACCAAGC	6055
Oy	1455	LysLeuLeuThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThr	1474
Dd	6056	AAACCTCCGACTTCTTCAACCAAGTACTGTACCAACCAACCAATTTACCACTAGAAACA	6115
Oy	1475	ValAlaLeuProValThrProValValThrProGlyGlnValSerThrProValSerThr	1494
Dd	6116	GTTGCTTTGCCAGTTACACAGAGGTGCACACTGAGCAAGTTTGTACCCACAGTGACACT	6175
Oy	1495	ThrThrSerGlyValLysProGlyThrThrAlaProSerLysProProLeuThrThrAlaPro	1514
Dd	6176	ACTACATACAGAGATGAACCTGGAACTGCTCCCTCCAAAGCCACTCTTAACTAAGGCTTCGG	6235
Oy	1515	ValLeuProValGlyThrGlnLeuProAlaGlyThrLeuProSerGlnGlnLeuProPro	1534
Dd	6236	GTGTCGCCAGTGGGTACTGAACCTTCAGCAAGTACTCAACCAACGAGAGCTCCACCT	6295
Oy	1535	PheProGlyProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeu	1554
Dd	6296	TTTCCAGAACCTTCTCTTAAACCAAGTCCACGACCTCTAAGAGATCTTGAATGCTCAATTG	6355
Oy	1555	ArgArgThrLeuSerProGlu***IleThrValThrSerAlaValGlyProValSerMet	1574
Dd	6356	AGAGAAACACTTAGTCCAGAGATATATACAGTGACTTCTGGGGTGGTCTGTGTCCATGG	6415
Oy	1575	AlaAlaProThrAlaIleThrGlnAlaGlyThrGlnProGlnLysGlyValSerGlnVal	1594
Dd	6416	GCGGCTCAACAGCAATACACGAAGACGAAGAAACAAGCTCTGAGAGGGGGTTTCCAGTGC	6475
Oy	1595	LysGlnGlyProValIleLeuAlaThrSerSerGlyAlaGlyValIleLysMetGlyArgPhe	1614
Dd	6476	AAAAGAGCCCTGTGCTTAGCAACTGATGCAGGAACTGATGTTTTTAAAGATGGGACGATTTT	6535

QY	1615	GLNVLSERVALAIAAASGLVALAGLNLVSGVLVYVAASLVSESLVSPAL	163
Db	6536	CAGGTTTCTGTTGCAGCAGACGGTCCCGAAGAGGGGAAAAAATGACCAAGAATGCA	659
QY	1635	LYSSEVALHISPEGLUSERSETHSERGIUSERSEVALLEUSERSESEPRO	165
Db	6596	AAGTCTGTTCAATTGGATTCAGACCTCAGAGTCTCTAGTCTCTCAAGTAGTAGTCCA	665
QY	1655	GLUSERTHIRLEUVALLYEPROGLUPROANGLYILETHRIIEPROGLVLESESERASP	167
Db	6656	GAGAGTACCTTGSTGTAACACAGAGCCGATGGCAATACCATCCCTGATCTCTTCAAGAT	671
QY	1675	VALPROGLUSERALAHISLYSTHTRHRLASERGVNALVYSESERASPETHGLVGLINPRO	169
Db	6716	GTGCCAGAGAGTGCCCAAAAACACTACTGCTCAGAGGCCAAAGTCAAGCACTGGGGAGCTT	677
QY	1695	THIRYVALGLYARXPEGLHVALTHRTHTRHRLAAENLYSVALGYARXPHESERVAL	171
Db	6776	ACCAAGTTTGGACGTTTTCAGGTGCACTACAGCAACCAAGTGGTGGTTTCTCTGTA	683
QY	1715	SELYSTHRLVAPPLYELLETHASPHRLYSLSGLVGLYPROVALHISERPROPRO	173
Db	6836	TCAAAACCTGAGGACAAGATCATCTGACCAAGAAAGAGAACCATGTGGCATCTCTCTCT	689
QY	1735	PHEMEAPLEUGLUGLINALVALLEUPROALVALIIEPROLYSGLVLYSPROGLU	175
Db	6896	TTTATGATTTGGAAACAAGCTGTTCTCTCTGCTGTGATACCAAGAAAGAAAGAGCTTGA	695
QY	1755	LEUSERGLUPROSETHISLEUANGLYPROSESERASPPROGLUALALAPHELEUSER	177
Db	6956	CTGTCAAGGCTTCAACATCTAAATGGGCGCTTCTGTGACCGAGGGCGCTTTTAAAGT	701
QY	1775	ARGASPVLAASAPSGLYSERGLVSESPROHISERPROHISGLINLEUSERSELYSESR	179
Db	7016	AGGATGTGGATGATGATGTTCCGGTAGTCCACACTGCCCCATCACTGAGCTCAAGAGCC	707
QY	1795	LEUPROSEGLINASINLEUSERGLINSEUSERANSERPHEAMSESESLYTWETSESR	181
Db	7076	CTTCTACCCAGAACTCTAAAGTCAAGCCTTAGTATTCATTAACTCTCTTACATGAGT	713
QY	1815	SERASPAENGLUSERASPLIEGLVJASPGVLSUPLEULYSLEUGLILEUARGLEUNG	183
Db	7136	AGCCAGCATGTGTCAGATATGAAAGATGAAACCTTAAATTAGAGCTCGACGACTACGA	719
QY	1835	ASPLYSHISLEULYSGULIEGLINASPLEUGLINSEARGLINLYSHISGLILIEGLUSER	185
Db	7196	GATAAACAATCTCAAGAAGATTCAAGACCTCGAGAGTGGCCAGAAAGCATGAAATTGAATCT	725
QY	1855	LEULYRTHIRLYSLEUGLYLYSVALPROPROALVALIIELEPROPROALALAPROLEU	187
Db	7256	TTGATATCCAAACGTGGCGAAGTCCCTCTGCTTATTATTTCCCGAGCTGCTCCCTT	731
QY	1875	SERGLYARGLARGLARGLPROTHIRYSESLYSGLYSEIRYSESESRARISSESRSESR	189
Db	7316	TCAGAGAGAAACACACACACCACTMAAAGCAAGCAAGCAAAATTAAGTCAGACAGATTCC	737
QY	1895	LEUGLYASNLYSESRPROGLINLEUSERGLVASNLEUSERGLYGLINSERALALASERVAL	191
Db	7376	TTGGGAGATAAAGACCCCGAGCTTTCAAGTAACTGTCTGTGCTCAGCTGACAGCTTCAGCT	743
QY	1915	LEUHSIPROGLINLINTHLEUHSIPROPROGLVASNLIIEPROGLUSERGLVGLINASGLIN	193
Db	7436	TTTGACCCCCACGACGACCTTCCACCTCTCTGGCAACATCCAGAGTCCGGGCAAAATCAG	749
QY	1935	LEULEUGINPROLEULYSPROSESRPROSESESRASPNLEULYSESERALAPHETHSESR	195
Db	7496	CTGTTACAGCCCCCTTAAGCCATCTCCCTCCAGTGACAACCTATTCAGGCTTCAACAGT	755
QY	1955	ASPGLYALALIESERVALPROSELEUSERALAPROGLYGLINLYTHIRSESESRTHASN	197
Db	7556	GATGTGTCCATTTAGTACCAAGCTTCTGTGCTCCAGGTCAAGGTAAATAAGCAACATC	761

QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAla 1985
 Db 7616 ATCGTCCAAAAACATATAATGAGATGTTGCC 7648

RESULT 13
 US-10-767-471-449
 ; Sequence 449, Application US/10767471
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: C1001505
 ; CURRENT APPLICATION NUMBER: US/10/767,471
 ; NUMBER OF SEQ. ID NOS: 50231
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ. ID NO 449
 ; LENGTH: 9417
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(9417)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
 US-10-767-471-449)

Alignment Scores:
 Pred. No.: 0 Length: 9417
 Score: 9737.00 Matches: 1963
 Percent Similarity: 88.03% Conservative: 1
 Best Local Similarity: 87.99% Mismatches: 21
 Query Match: 90.06% Indels: 246
 Gaps: 1

US-10-010-720-14 (1-2136) x US-10-767-471-449 (1-9417)

QY 1 MetSerGlyValAlaGlnValGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 956 ATGTCTGGCGCGCCCGAGAGACAGACAGACACTCCCGTTCCCTGCTCTCGCGC 1015

QY 21 ProAlaProAlaProIlyAsnGlySerSerSerSerSerSerValGlyIlyLeuGly 40
 Db 1016 CCGGCTCTCCCGCCCAAGATGCTCCAGCTCCGATCTCCCGGGGGAGAACTGGGA 1075

QY 41 AlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 60
 Db 1076 GCGCGCGCGCGCGCGCTGTGACCGGACGAGACGAGAGTACAGCGCGCGCCACACT 1135

QY 61 MetAspIlyAspSerSerArgGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 80
 Db 1136 ATGAGCAAGACAGACCGCTGGGGCGCGCGACCACTACCACTAGACACCGCTTCTTC 1195

QY 81 ArgArgSerValIleCysAspSerSerAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 100
 Db 1196 CGCGGAGCTCATCTGNGACTCCCAATGCCACTGACCTGAGCTTCCGGGCTTCTCTT 1255

QY 101 SerLeuProGlnProSerIleProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 120
 Db 1256 TCCCTGCGCGACCGACATCCCGCGCTGTCCCGAGAGTGTCCACCGAGACCGCCAC 1315

QY 121 ArgGlnIleThrValThrAlaThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 140
 Db 1316 CGGGAAGAGCCGTGACCGCGCACCTTCCAGGTAGCCACGAGCGCTCCAGCGCT 1375

QY 141 AlaAlaProGlyGlnIleAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 160
 Db 1376 RCGCGCCCTGGGAAACAGCGCTGCGCGGCGCTGCGCCCTGCACTGTCCACAGATACC 1435

QY 161 SerIlyAspArgProValSerGlnProSerLeuValGlySerIlyGlnIleProPro 180
 Db 1436 AGCAAAACCGCCAGTGTCCAGCTTGTGTGGAGAGCAAAAGAGAGCGCGCGC 1495

QY 181 AlaArgSerGlySerGlyGlySerAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 200

Db 1496 GCGAGAAATGGCAGCGCGCGCGCACGCGCAAGAGACCAAGAGAGACGAGACCGACG 1555

QY 201 GlnAspAspIleGlnIleGlnIleThrValAlaAlaAlaAlaAlaAlaAlaAlaAla 220
 Db 1556 CAGGATGATATCGAAGAGCTGGAGCAAGCGCGGTGGAAATGTCTAACGATGCGCTT 1615

QY 221 LeuIlyPheAspIleGlnIleGlyArgGlySerPheIleThrValIleIlyGlyLeuAsp 240
 Db 1616 CTCAGTTTGAATCGAATTCGCAAGGCTCTTTAAGCGGTCTACCAAGTCTGAC 1675

QY 241 ThrGlnIleThrValGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 260
 Db 1676 ACTGAACACCGCTGGAAGTCCGCTGTGTGAATCGAGATGAAATTAACAAAGTCT 1735

QY 261 GlnArgGlnArgPheIlyGlnIleAlaGlnIleLeuIlyGlyIleGlnIleProAsnIle 280
 Db 1736 GAGAGCGAGAGATTTAAAGAGAGTGAATGTAAAGTCTTACGACATCCCAATATT 1795

QY 281 ValArgPheTyrAspSerTyrGlnSerThrValIlyGlyIlyIlyCysIleValIleVal 300
 Db 1796 GTTAGATTTTATATCTCTGGAGATCCAGTAAAGAGAGATGATGTTTGTG 1855

QY 301 ThrGlnIleMetThrSerGlyThrIleIlySerThrTyrIleIlySerPheIlyValMetIly 320
 Db 1856 ACTGAACCTATGACGCTGGACACTTAAACGATCTGAAAGGTTTAAAGTATGAG 1915

QY 321 IleIlyValIleuArgSerTyrPyrGlnIleIleuIlyGlyIleuIlePheIleuIle 340
 Db 1916 ATCAAGTTTCTAAGAGCTGGTGCCTGCAGATCTTAAAGTCTTACGTTCTTCACT 1975

QY 341 ArgThrProProIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 360
 Db 1976 CGAACTCCACTTATCATCTTACCGCGCATCTTAATGTGACACATCTTATCACCGGCT 2035

QY 361 ThrGlySerValIlyIleGlyAspLeuGlyLeuAlaThrIleuIlySerAlaSerPheAla 380
 Db 2036 ACTGGCTCAGTCAAGATGAGACCTGCTGTGCAACCTGAAGCGGCTTCTTTGGC 2095

QY 381 IlySerValIleIleIlyThrProGlnIlePheMetAlaProGlnIleIlyGlnIlyIly 400
 Db 2096 AAGAGTGTATAGTATCCCGAGAGTCTATGCGCTGAGATGATAGAGAAATATGAT 2155

QY 401 GluSerValAspValIlyAlaPheGlyMetCysMetLeuGlnMetAlaThrSerGlnIly 420
 Db 2156 GAATCCGTGACCTTATGCTTTGGAGTGTGATGCTTGAATGCTATCATCTGAATAT 2215

QY 421 ProIlySerGlnCysGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 440
 Db 2216 CTTACTCGGAGTGCAGAAATGCTGCRAGATCTACCGTGGCTGACCACTGGGGTGAAG 2275

QY 441 ProAlaSerPheAspIlyValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 460
 Db 2276 CCAGCCAGTTTACCAAGATGACATTCCTGAAGTGAAGAAATATATGAAGATGACATA 2335

QY 461 ArgGlnAsnIlyAspGlnArgIlySerIleIlyAspLeuIleuIleAlaPheGln 480
 Db 2336 CGACAAACCAAGATGAATATATTCATCAAGACCTTTGAACATGCTTTCCTCA 2395

QY 481 GlnGlnIleThrIlyAlaArgValGlnLeuAlaGlnIleuAspAspGlyIlyIleAlaIle 500
 Db 2396 GAGGAAACAGAGATGACGGGTAGATTTAGCAGAAAGAGATGAGAGAAAAAATGACATA 2455

QY 501 IlyLeuIleThrLeuArgIleGlnAspIleIlyValIleIlyIlyIlyIlyIlyIlyIly 520
 Db 2456 AAATTTATGCTAGATTAAGATATTAAGAAATTAAGAGGAAATTAACAAAGATATGAA 2515

QY 521 AlaIleGlnPheSerPheAspLeuGlnArgAspValProGlnIlyValAlaGlnIleu 540
 Db 2516 GCTATTGAGTTTCTTTGATTTAGAGAGAGATGCTCCAAAGATGTGCAAGAAATG 2575

QY 541 ValGlnSerGlyIlyThrValCysGlnIlyAspIleIlyThrMetAlaIlyAlaIleIlyAsp 560

Db 2576 GTAGAGTCTGGGTATGTCTGTGAAGGTGATCAAGACCATGGCTAAAGCTATCAAAAGAC 2635
QY 561 ArgValSerLeuIleuValArgLysArgLysGlnLeuValArgGlnGlnGln 580
Db 2636 AGAGTATCATTAATTAAGAGAAACGAGAGCGGAGTTGTGACGGAGAGCAAGAA 2695
QY 581 LysLysLysGlnGlnGlnSerSerLeuLysGlnGlnValGlnGlnSerSerLysGln 600
Db 2696 AAAAAAGAGAGAGAGAGAGAGTCTCAACAGCAGGTAGAACAAATCAGTCTCCAG 2755
QY 601 ThrGlyIleLysGlnLeuProSerAlaSerThrGlyIleProThrAlaSerThrThrSer 620
Db 2756 ACAGGAATCAAGAGCTCCCTCTGTGACACGGGACCTACTGCTTCTACCACTTCA 2815
QY 621 AlaSerValSerThrGlnValGlnProGlnGlnProGlnLysAspGlnHisGlnGlnLeu 640
Db 2816 GCTTCAGTTTCTACACAGTAGAAGCTGAGAGAACTGAGGCAATCAACTCAACACTCA 2875
QY 641 GlnTyrGlnGlnProSerLysSerValLeuSerAspGlyThrValAspSerGlyGlnGly 660
Db 2876 CAGTACAGCAGACCCAGTATATCTGTATCTGATGGAGCGGTTGACAGTGGTCAAGGA 2935
QY 661 SerSerValPheThrGlnSerArgValSerSerGlnGlnThrValSerTyrGlySerGln 680
Db 2936 TCTCTGTCTTCAAGATCTCAGATGAGAGCCAAACAGCAGTTTCATATGGTTCCAA 2995
QY 681 HisGlnGlnAlaHisSerThrGlyThrValProGlnHisIleProSerThrValGlnAla 700
Db 2996 CATGAACAGGCACATTTCTACAGGACAGTCCAGGCGATATACCTTCTACTGTCACAGCA 3055
QY 701 GlnSerGlnProHisGlyValTyrProProSerSerValAlaGlnGlnGlnSerGlnGly 720
Db 3056 CAGTCTCAGCCCATGGGTATATCCACCTCAAGTGGCACAGGGGACAGGACAGGT 3115
QY 721 GlnProSerSerSerSerLeuThrGlnValSerSerSerGlnProIleGlnHisProGln 740
Db 3116 CAGCATCTTCAAGTATCTTAACAGGGGTTTCATCTTCCCAACCCATCAACAACTCTCAG 3175
QY 741 GlnGlnGlnGlnLysIleGlnGlnThrAlaProProGlnGlnThrValGlnTyrSerLeuSer 760
Db 3176 CAGCAGAGGAGATACAGCAGACAGCCCTCTCAACAGCAGTTCATTTCACTTCA 3235
QY 761 GlnThrSerThrSerSerGlnAlaThrThrAlaGlnProValSerGlnProGlnAlaPro 780
Db 3236 CAGACATCAACCTCCAGTGGAGCCACTACTGCAAGCAGCAGTGAAGTCAACTCAAGCTCCA 3295
QY 781 GlnValLeuProGlnValSerAlaGlyLysGln----- 791
Db 3296 CAAAGTCTGCTCAAGATATCAGCTGGAAAAACAGCTTCCAGTTCCAGCCAGTACCAACT 3355
QY 791 ----- 791
Db 3356 ATCCAGAGGGAACCTCAGATCCAGTTGGACACAAACCCCTCGGTTGTTCAGTCCACTCT 3415
QY 791 ----- 791
Db 3416 GGTGCTCATTTCTTCCAGTGGAGACAGCGCTCCCTACTCCTTGTCTCCTCAGTACCT 3475
QY 791 ----- 791
Db 3476 GTCTTCARATTTCCCATATCACTCCTCATGTGTCTACGGGTGACAGAGGTTTCTCATCC 3535
QY 791 ----- 791
Db 3536 CTTCCCATCAATATGGAGAGTGGCATTTACTCAGCCTCTGTCTACGTTGGCTTATCTGCT 3595
QY 791 ----- 791
Db 3596 ACAACAGCTGGCATCCCGGGGTATCAACTGTGGTTCTTAGTACAGTTCCAACTTTCTG 3655
QY 791 ----- 791
Db 3656 CAGCCTGTAGTACGCTGCCAAGTCAGGTTTCACCCAAGCTCCTACAAACGAGATTGAC 3715

QY 791 ----- 791
Db 3716 TCATGGGAATTACAGCTAACTTGGACAAAGCTGTGAGTTCCACTTCTTGAGAT 3775
QY 791 ----- 791
Db 3776 GTTCTGTACAGGGCTTCCACTCGACTGACCTGCACACAGTACCCAGAGATTCAAATATT 3835
QY 791 ----- 791
Db 3836 GCTCCCTCTTCAACGTGGCTGTGTTGATCATTCATTAAGTCTTATCCCTCCATG 3895
QY 791 ----- 791
Db 3896 CCGACAGAGTACTGGCTACCTGGGTACTTCCACAGTGTGACGCTTATGTGAA 3955
QY 791 ----- 791
Db 3956 TCAATCTTTAGTCTCATGGTGTGTAGAGAGCAGGTTCAAGTCCAGCCAGGA 4015
QY 792 ----- 792
Db 4016 GGGAGTTTACACAGACCCCACTACATCTCCAGCAGAGTTTGGAGAGTACTCAG 4075
QY 795 GlyValSerGlnValAlaProAlaGlnProValAlaValAlaGlnProGlnAlaThrGln 814
Db 4076 GAGTCTCTCAGGTTGCTCTGTCAGAGCCAGTTGTCAGTACAGCAAGCCAACTACCCAG 4135
QY 815 ProThrThrLeuAlaSerSerValAspSerAlaHisSerAspValAlaSerGlyMetSer 834
Db 4136 CCGACCACTTGGCTTCTCTGTAGACAGTGCACATTCAGATGTGGTTCAGGATATAGT 4195
QY 835 AspGlyAsnGlnAsnValProSerSerSerGlyArgHisGlnGlnGlnGlnThrThrLysArg 854
Db 4196 GATGGCAATGAGAACCTCCATCTTCCAGTGAAGGACATGAAAGAACTTCAAAAACGG 4255
QY 855 HisTyrArgLysSerValArgSerArgSerArgHisGlnLysThrSerArgProLysLeu 874
Db 4256 CATTACGAAAAATCTTAAGAGACTCCCTCCGACATGAAAAAATCTTACGCCCAAAATTA 4315
QY 875 ArgIleLeuAsnValSerAsnLysGlyAspArgValAlaGlnCysGlnLeuGlnThrHis 894
Db 4316 AGAATTTTGAATGTTTCAATAAAGAGACCGAGTAGTGAATGATTAAGAGACTCAT 4375
QY 895 AsnArgLysMetValThrPheLysPheAspLeuAspGlyAspAsnProGlnGlnIleAla 914
Db 4376 AATAGAAAAATGGTTACATTCAATTGACCTTAGATGTGACAAACCCGAGGAGATAGCA 4435
QY 915 ThrIleMetValAsnAsnAspPheIleLeuAlaIleGlnLysGlnSerPheValAspGln 934
Db 4436 ACAATTTATGTAACAATGACTTATTTCTAGCAATGAGAGAGAGTCTTGTGATCA 4495
QY 935 ValArgGlnIleIleGlnLysAlaAspGlnMetLeuSerGlnAspValSerValGlnPro 954
Db 4496 GTGCGAGAAATTAATGAAAAAGCTGTAAGAAATCTCAGTAGAGATGTGAGAAACCA 4555
QY 955 GlnGlyAspGlnGlnLysLeuGlnSerLeuGlnGlnLysAspAspTyrGlyPheSerGlySer 974
Db 4556 GAGGGTGATCAGGAGATTGAGAGTCTACAGGAAAAAGATGACTATGGCTTTTGAAGTCT 4615
QY 975 GlnLysLeuGlnGlnLysLysGlnProIleProAlaSerSerMetProGlnGlnIle 994
Db 4616 CAGAAATTTGAGAGAGAGTTCACAAACAACCAATTCCTCGCTTCCATGCACAGCAATA 4675
QY 995 GlyIleProThrSerSerLeuThrGlnValAlaHisSerAlaGlyArgArgPheIleVal 1014
Db 4676 GGCATTTCTTACCAAGTCTTAACTCAAGTTGTTCATTTGCGGGAAGCGGTTTATAGTG 4735
QY 1015 SerProValProGlnSerArgLeuArgGlnSerLysValPheProSerGlnIleThrAsp 1034
Db 4736 AGTCTGTCCAGAAAAGCCGATTACGAGATCAAAAGATTTTCCCGCAGTAAATPAAGAT 4795

Db 6956 CTGTCAGAGCCTTCACTTAATATGGCCGCTTTCTGACCCGAGGCCGCTTTTAAAGT 7015
 1775 ArgAspValAspAspGlySerGlySerProHisSerProHisGlnLeuSerSerLeuSer 1794
 7016 AGGATGTGATGTGTTCCGGTATCCACACTCGGCCCATCGCTGAGCTCAAAAGC 7075
 QY 1795 LeuProSerGlnLeuLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSer 1814
 7076 CTTCCTAGCCAGATCTTAAGTCAAAAGCCTTAGTATTCATTAACTCTTACATGAGT 7135
 Db 1815 SerAspAsnGlnSerAspGlnGlnAspGlnAspLeuLeuLeuGlnGlnGlnGln 1834
 7136 AGGACATATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7195
 QY 1835 AspHisSerLeuSerGlnGlnGlnAspLeuGlnSerGlnGlnGlnGlnGlnGlnGln 1854
 7196 GATTAACATCTCAAAAGATTTCAAGACTTCAGAGCTTCAGAGCTTCAGAGCTTCAGAG 7255
 Db 1855 LeuTyThrIleuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1874
 7256 TTGTATACCAAACTGGGCAAGTGGCCCTGCTGTATATTCCTCCAGCTGCTCCCTT 7315
 QY 1875 SerGlyArgArgArgArgArgArgArgArgArgArgArgArgArgArgArgArgArg 1894
 7316 TCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7375
 QY 1895 LeuGlyAsnLeuSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerVal 1914
 7376 TTGGGGAATAAAGCCGCCGCTTTCAGGTATCTGCTGCTGCTGCTGCTGCTGCTGCT 7435
 Db 1915 LeuHisProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1934
 7436 TTGACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7495
 QY 1935 LeuLeuGlnProLeuLeuSerProSerProSerSerAspAsnLeuTySerAlaPheThrSer 1954
 7496 CTGTTCAGCCCTTCAAGCATCTCTCCAGTACCACTCTTAATTCAGCTTCAAGCAT 7555
 Db 1955 AspGlyAlaIleSerValProSerLeuSerAlaProGlyGlnGlnGlnGlnGlnGln 1974
 7556 GATGGTGCATTTTCAGTACCAAGCTTTCTGCTCCAGGTCAAGGTATTAAGACCAATC 7615
 QY 1975 ThrValGlyAlaThrValAsnSerGlnAlaAla 1985
 Db 7616 ATGTCCTCAAAACATTAATGAGATGTGCC 7648
 RESULT 14
 US-09-979-167-63
 / Sequence 63, Application US/09979167
 / GENERAL INFORMATION:
 / APPLICANT: PLOMAN, GREGORY D.
 / APPLICANT: MARTINEZ, RICARDO
 / APPLICANT: WHYTE, DAVID
 / APPLICANT: SUDERSANAM, SUCHA
 / TITLE OF INVENTION: PROTEIN KINASES
 / FILE REFERENCE: 038602/1273
 / CURRENT APPLICATION NUMBER: US/09/979,167
 / NUMBER OF SEQ ID NOS: 269
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 63
 / LENGTH: 7328
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: modified base
 / LOCATION: (2083)..(2084)
 / OTHER INFORMATION: a, t, c, g, other or unknown
 / US-09-979-167-63
 Alignment Scores:
 Pred. No.: 0 Length: 7328
 Score: 9729.50 Matches: 1939

Percent Similarity: 97.68% Conservative: 1
 Best Local Similarity: 97.63% Mismatches: 17
 Query Match: 89.99% Indels: 29
 DB: 5 Gaps: 2
 US-10-010-720-14 (1-2136) x US-09-979-167-63 (1-7328)
 QY 1 MetSerGlyGlyAlaAlaGlnGlnGlnSerSerThrProGlySerLeuPheLeuSerPro 20
 Db 42 ATGTTCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 101
 QY 21 ProAlaProAlaProLeuSerGlnGlnSerSerSerSerSerSerSerValGlyGlnGlnGly 40
 Db 102 CCGGCTCTGCCCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 161
 QY 41 AlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 60
 Db 162 GCCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 221
 QY 61 MetAspLysAspSerArgGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 80
 Db 222 ATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 281
 QY 81 ArgArgSerValIleCysAspSerSerSerSerSerSerSerSerSerSerSerSerSer 100
 Db 282 CGCCGAGCGTCATCTGCGACTCCAAATGCACTGCACTGAGAGCTTCCGCGCTTCCCTT 341
 QY 101 SerLeuProGlnProSerSerLeuProAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 120
 Db 342 TCCCTCCCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 401
 QY 121 ArgGlnGlnThrValThrAlaThrAlaThrSerGlnValAlaGlnGlnProProAlaAla 140
 Db 402 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 461
 QY 141 AlaAlaProGlyGlnGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 160
 Db 462 GCCGCCCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 521
 QY 161 SerLysAspArgProValSerGlnProSerLeuValGlySerLysGlnGlnProPro 180
 Db 522 AGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 581
 QY 181 AlaArgSerGlySerGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
 Db 582 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 641
 QY 201 GlnAspAspIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
 Db 642 CAGATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 701
 QY 221 LeuLysPheAspIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
 Db 702 CTCGAATTTGACATCCAAATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 761
 QY 241 ThrGlnThrThrValGlnValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 260
 Db 762 ACTGAACACAGCGTGAAGTGGCTGTGTGAATCTCAAGATCGAATAATTAACAAAGTCT 821
 QY 261 GlnArgGlnArgPheLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 Db 822 GAGAGGACAGATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 881
 QY 281 ValArgPheThrAspSerThrProSerThrValValGlyLysLysCysIleValLeuVal 300
 Db 882 GTTAGATTTATCATCTCTGGAGATCCACAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 941
 QY 301 ThrGlnLeuMetThrSerGlyThrLeuLysThrTyLeuLysArgPheLysValMetLys 320
 Db 942 ACTGAACCTATGACGCTGGAACACTTAAGAGTATCTGAAGAGAGAGAGAGAGAGAGAG 1001
 QY 321 IleLysValLeuArgSerThrProCysArgGlnIleLeuLysGlyLeuGlnPheLeuHisThr 340

QY 1060 AlaPheSerGluLeuAlaGatGAlaGlnMetThrGluGlyProAsnThrAlaProProAsn 1079
DB 3138 GCCCTTTCCTAACTAGACGTCGCCAATATGACAGAAAGACCCAAATATACAGACCTCCCAAC 3197
QY 1080 PheSerHisThrGlyProThrPheProValValProPheLeuSerSerIleAlaGly 1099
DB 3198 TTATGATATACAGAACCAACATTTCCAGTAGTACTCTCTTCTTAATATACATTGCTGGA 3257
QY 1100 ValProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProProAsnAsp 1119
DB 3258 GTCCCAACCAAGAGAGACCAAGCAAGCAAGTCCTCTCCCAACAGACCCCTCTCAATAGAC 3317
QY 1120 IleSerHisSerValIleGlnSerGluValThrValProThrGluGluGlyIleAlaGly 1139
DB 3318 ATTTCCATCATGAAATTCAGTCTGAGGTACAGTGCCCACTGAAGAGGGGATTGCTGGA 3377
QY 1140 ValAlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProValSerGlu 1159
DB 3378 GTTGCCACCAAGCAAGGTGTGTAACTTCAAGTGCTCTCCCAATACCACTGTGTGAA 3437
QY 1160 SerProValLeuSerSerValValSerSerIleThrIleProAlaValIleSerIleSer 1179
DB 3438 TCACCAAGTACTTCCAGCGTAGTTTCAAGATACAAATACCTGCAAGTGTCTCAATATCT 3497
QY 1180 ThrThrSerProSerLeuGlnValProThrSerThrSerGluIleValValSerSerThr 1199
DB 3498 ACTACATCCCCGTCACCTTCAAGTCCCAATCCCACTGAGATCGTTGTTTCTAGTACA 3557
QY 1200 AlaLeuThrProSerValThrValSerAlaThrSerIleSerAlaGlyGlySerThrAla 1219
DB 3558 GCACGTATCTTCAGTAAAGTTTCAAGCACTTCAAGCTCTGAGGGGAGACTGCT 3617
QY 1220 ThrProGlyProLysProProAlaValValSerGlnIleAlaGlySerThrThrVal 1239
DB 3618 ACCCAAGTCTTCAAGCTCCAGCTAGTATCTCAGAGGACAGGACACACATCTAGTGTG 3677
QY 1240 GlyAlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSerGlnLeu 1259
DB 3678 GGAAGCCATTTAATCATGATTTCTACCAACCACTTCCCAAGCAAGCTTCCAGACTG 3737
QY 1260 SerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlnThrValValValSer 1279
DB 3738 TCCATTAGCTTACAGACAGACTTCTTACCTCACTTAACTGAACCGGTAGTATAGC 3797
QY 1280 AlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSerLeuSer 1299
DB 3798 GCACACTCACAATGATTAAGACATCTCATAGCAAGTACAACTGGATTGCTTCTCCCTCT 3857
QY 1300 AlaProSerSerSerSerSerProGlyAlaGlyValSerSerTyrlleSerGlnProGly 1319
DB 3858 GCACCAATCTTCCCTCTCTCTCTGAGAGAGGTCTAGTTATATTCTCAGCCGTGT 3917
QY 1320 GlyLeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuProGlnAla 1339
DB 3918 GGGCTGATCTTGGTCATTCATCATGATAGCTTCACTCTTCTTCCCAAGCA 3977
QY 1340 AlaGlyProThrSerThrProLeuLeuProGlnValProSerIleProProLeuValGln 1359
DB 3978 GCAGAGACTTCTTACACTTTATTTACCCCAAGTACTTGTATCCACCTGTTGGTAG 4037
QY 1360 ProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGlnProAla 1379
DB 4038 CCTGTGGCAATGTGCTGCTGTACAGCAACACTAATTCATAGTCAGCTCAACAGCT 4097
QY 1380 LeuLeuProAsnGlnProHisThrHisCysProGluValAlaAspSerAspThrGlnProLys 1399
DB 4098 TTACTTCCCAACAGCCCACTACTCATTTGCTGAATATGATTTCTGATTAACAACCCAAA 4157
QY 1400 AlaProGlyIleAspAspIleLysThrLeuGluGluLysLeuArgSerLeuPheSerGlu 1419
DB 4158 GCTCTCGAATTGATGACATTAAGACCTGTAAGAAAGAAAGTGGGTCTCTGTTCAAGTAA 4217
QY 1420 HisSerSerSerGlyAlaGlnHisAlaSerValSerLeuGluThrSerLeuValIleGln 1439

DB 4218 CACAGCTCATCTGAGACTGAGACTGCTGTCTCACTGAGAGACTCATAGTATAGAG 4277
QY 1440 SerThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeuThrSer 1459
DB 4278 AGCAGTGTACACAGGACATCCCACTACTGCTGTGGACCAAGCAAACTCTGACTTCT 4337
QY 1460 ThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeuProVal 1479
DB 4338 ACCACAAGTACTTGTCTTACCAACCAATTTACCTAGAAAGAGTGTGTTGGCCGTT 4397
QY 1480 ThrProValValThrProGlyGlnValSerThrProValSerThrThrThrSerGlyVal 1499
DB 4398 ACACCAAGTGTGACACTGGGCAAGTTTCTAACCCAGTCCAGACTACTACATCAGAGAGT 4457
QY 1500 LysProGlyIleThrAlaProSerLysProProLeuThrLysAlaProValLeuProValGly 1519
DB 4458 AAACCTGGAACCTCTCTCCCAAGCACTTAACTAAGGCTCCGGTGTGCCAGTGGGT 4517
QY 1520 ThrGluLeuProAlaGlyThrLeuProSerGluGlnLeuProPheProGlyProSer 1539
DB 4518 ACTGAATTCACAGAGTACTTACCCAGAGACAGCTGCCACTTTTCCAGAGACTTCT 4577
QY 1540 LeuThrGlnSerGlnProLeuGluAspLeuAspAlaGlnLeuArgArgThrLeuSer 1559
DB 4578 CTAAACCAGTCCCAAGCAACCTTAGAGAGATCTTGAATGCTCAATTGAGAAAGAACCTTAG 4637
QY 1560 ProGlu**IleThrAlaThrSerAlaValAlaGlyProValSerMetAlaAlaProThrAla 1579
DB 4638 CACGAATGATCAACAGTACTTGGGTGGTGTGGTGTGCTCAAGGGGCTCCAAACAGCA 4697
QY 1580 IleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGluGlyProVal 1599
DB 4698 ATCACAAGAGCAAGAAACAAGCTTCAAGAGGAGTTTCTCAATCAAGAAGGCCCTGTC 4757
QY 1600 LeuAlaThrSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSerValAla 1619
DB 4758 CTAGCAACTGATTCAAGAGTGTGTTTAAATAGGAGCAATTCAGGTTCTGTTGCA 4817
QY 1620 AlaAspGlyAlaGlnLysGluGlyLysAsnLysSerGluAspAlaLysSerValHisPhe 1639
DB 4818 GCAGAGGTGCCCCAAGAAAGGGTAAATAATTAATCAAGATCAAAAGTCTGTTCATT 4877
QY 1640 GluSerSerThrSerGluSerSerValLeuSerSerSerSerProGluSerThrLeuVal 1659
DB 4878 GAATCCAGACCTCACAAGTCTCAGTGTATCAAGTAGTAGTCCAGAGATCTTGCTG 4937
QY 1660 LysProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGluSerAla 1679
DB 4938 AAACCAAGGCCGAATGGCATTAACATCCCTGATATCTTCAAGATGTCAGAGAGTGC 4997
QY 1680 HisLysThrThrAlaSerGluAlaLysSerAspThrGlyGlnProThrLysValGlyArg 1699
DB 4998 CACAAAACCTACGTCTCAGAGGCAAGTCAAGACATGGGAGCTTCAAGAGTTGACGT 5057
QY 1700 PheGlnValThrThrThrAlaLeuLysValGlyArgPheSerValSerLysThrGluAsp 1719
DB 5058 TTTCAGGTGACAATCAGAAACAAAGTGGTCTTCTCTGTATCAAAAACAGAGAC 5117
QY 1720 LysIleThrAspThrLysLysGluGlyProValAlaSerProProPheMetAspLeuGlu 1739
DB 5118 AAGATCACTGACACAAGAAAGAAAGACAGTGGCATCTCTCTTTTATGATTGGAA 5177
QY 1740 GlnAlaValLeuProAlaValIleProLysLysGluLysProGluLeuSerGluProSer 1759
DB 5178 CAAGCTGTCTTCTCTGTGATTAACAAGAAAGAGCTTAAGTGTAGAGCTTCA 5237
QY 1760 HisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspValAspAsp 1779
DB 5238 CATCTAAATGGGCGCTCTTCTGACCCGAGAGCGCTTTTAAAGTAGGATGTGATGAT 5297
QY 1780 GlySerGlySerProHisSerProHisGlnLeuSerSerLysSerLeuProSerGlnAsn 1799

Qy	1038	AlaSerThrAlaGlnSerProGlyMetAsnLeuSerHisSerAlaSerSerLeuSerLeu	1057
Db	999	GCCCTTAAGCTCAGAGCCCTGGAAATGAACTGTCTCACTCGATCATCTCCCTTAAGCTTA	1058
Qy	1058	GlnGlnAlaPheSerGluLeuArgAlaGlnMetThrGlnGlyProAsnThrAlaPro	1077
Db	1059	CAACAGGCTTTTCTGAACTTAGAGCTGACGCTCCAAATGACAAAGAACCCANAAACGACCT	1118
Qy	1078	ProAsnPheSerHisThrGlyProThrPheProValValProPhePheLeuSerSerIle	1097
Db	1119	CCAAACTTAGTCAATACAGACCAACATTTCCAGTAAGTACTCTTCTTTTAAGTAGACTT	1178
Qy	1098	AlaGlyValProThrThrAlaAlaAlaThrAlaProValProAlaThrSerSerProPro	1117
Db	1179	GCTGAGATCCCAACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCT	1238
Qy	1118	AsnAspIleSerThrSerValIleGlnSerGlnValThrValProThrGlnGlnGlyIle	1137
Db	1239	AATGACATTTCCACATCAGTAAATTCAGTCTGAGGTACAGTGCCTGAAAGAGGAGAT	1298
Qy	1138	AlaGlyValAlaThrSerThrGlyValValThrSerGlyGlyLeuProIleProProVal	1157
Db	1299	GCTGAGATGCCACACAGCAGCAGGTTGGTACTTCAGGTGCTCCCCCATACACCTGTG	1358
Qy	1158	SerGluSerProValLeuSerSerValValSerSerIleThrIleProAlaValValSer	1177
Db	1359	TCTGAAATCACACAGTACTTCCAGCGTGAATTCAGATACAAATACCTGCACTGTCTCA	1418
Qy	1178	IleSerThrThrSerProSerLeuGlnValProThrSerThrSerGlnIleValValSer	1197
Db	1419	ATATCTACTACATCCCGTCACTTCAGATCCCAATCCACATCCAGATCGATGTTCT	1478
Qy	1198	SerThrAlaLeuThrProSerValThrValSerAlaThrSerAlaSerAlaGlySer	1217
Db	1479	AGTACACACAGTATCTCTTCAAGTAAAGTTTCAACAACTTCAGCTCTGACGGGGAGT	1538
Qy	1218	ThrAlaThrProGlyProGlyProProAlaValValSerGlnGlnAlaGlySerThr	1237
Db	1539	ACTGCTACCCAGGATCTTAAAGCTCCAGCTGATGATCTCAGCGAGCAGCAGCAGCAGT	1598
Qy	1238	ThrValGlyAlaThrLeuThrSerValSerThrThrThrSerPheProSerThrAlaSer	1257
Db	1599	ACTGTGGAGCCACATTAACATTCAGTTTCAACACACTTCATTCCTCCAAAGCAGCTTCA	1658
Qy	1258	GlnLeuSerIleGlnLeuSerSerSerThrSerThrProThrLeuAlaGlyThrValVal	1277
Db	1659	CAGCTGSCATTCAGTTAGCAGCACTTCTACTCTTACTTAAAGTAAACCTGTGTA	1718
Qy	1278	ValSerAlaHisSerLeuAspLysThrSerHisSerSerThrThrGlyLeuAlaPheSer	1297
Db	1719	GTTAGCCACACTCAGTATGAGATGATCTCATGAGCAGTACCACTGATGGCTTCTCC	1778
Qy	1298	LeuSerAlaProSerSerSerSerSerProGlyAlaGlyValSerSerThrIleSerGln	1317
Db	1779	CTCTCTGACCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1838
Qy	1318	ProGlyGlyLeuHisProLeuValIleProSerValIleAlaSerThrProIleLeuPro	1337
Db	1839	CTGTGGGGCTGATCTTGGTATTCATTCATTCATTCATTCATTCATTCATTCATTCAT	1898
Qy	1338	GlnAlaAlaGlyProThrThrThrProLeuLeuProGlnValProSerIleProProLeu	1357
Db	1899	CAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1958
Qy	1358	ValGlnProValAlaAsnValProAlaValGlnGlnThrLeuIleHisSerGlnProGln	1377
Db	1959	GTTACAGCTGTGGCAATGAGCTGCTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGC	2018
Qy	1378	ProAlaLeuLeuProAsnGlnProHisThrHisCysProGlnValAlaSerAspThrGln	1397
Db	2019	CCAGCTTTGGCTCCCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	2078
Qy	1398	ProLysAlaProGlyIleAspAspIleLysThrLeuGlnGlyLysLeuArgSerLeuPhe	1417
Db	2079	CCCAAGCTCTGGAATGATGACATGAAGCTCTAGAGAGAGAGAGAGAGAGAGAGAG	2138
Qy	1418	SerGlnHisSerSerSerGlyAlaGlnHisAlaSerAlaSerLeuGlnThrSerLeuVal	1437
Db	2139	AGTAAACAGCTTATCTGAGGCTCAGACATGCTCTCTCTCTCTCTCTCTCTCTCTCT	2198
Qy	1438	IleGluSerThrValThrProGlyIleProThrThrAlaValAlaProSerLysLeuLeu	1457
Db	2199	ATAGAGAGCATGTCACACAGCAGATCCCAACTACTGCTGTGTGACCAAGCAACTCTG	2258
Qy	1458	ThrSerThrThrSerThrCysLeuProProThrAsnLeuProLeuGlyThrValAlaLeu	1477
Db	2259	ACTTCTACCAAGTACTTCTGTTACCAACCAATTTACCATAGAGAACAGTGTCTTG	2318
Qy	1478	ProValThrProValValThrProGlyGlnValSerThrProValSerThrThrThrSer	1497
Db	2319	CCAGTTACACAGTGTGTCACCTGGGGAAGTTCTACCCACAGTCAACACTACTACATCA	2378
Qy	1498	GlyValLysProGlyThrAlaProSerLysProProLeuThrLysAlaProValLeuPro	1517
Db	2379	GGAGTGAACCTGSAACCTGCTCCCTCAAGCAGCAGCTTAAGCTCCGCTCTGCCA	2438
Qy	1518	ValGlyThrGluLeuProAlaGlyThrLeuProSerGlnGlnLeuProPheProGly	1537
Db	2439	GTGGGTACTGAACCTCCAGCAGGATCTACCCAGCAGCAGCTGCCACCTTTCCAGA	2498
Qy	1538	ProSerLeuThrGlnSerGlnGlnProLeuGlnAspLeuAspAlaGlnLeuArgThr	1557
Db	2499	CTTCTCTAACCCAGTCCACAGCACTCTGAGAGATCTTAAGTCTCAATGAGAAAGACA	2558
Qy	1558	LeuSerProGlu**IleThrValThrSerAlaValGlyProAlaSerMetAlaAlaPro	1577
Db	2559	CTTAGTCCAGAGATKATCAGATGATCTTGCGGTGTCTGTCTGTCTCATGCGGCTCYA	2618
Qy	1578	ThrAlaIleThrGluAlaGlyThrGlnProGlnLysGlyValSerGlnValLysGlyGly	1597
Db	2619	ACAGCAATCAACAGAGCAGGACACAGCCTCAGAGAGGTCTTCTCAAGCAAGAAAGGC	2678
Qy	1598	ProValLeuAlaThrSerSerSerGlyAlaGlyValPheLysMetGlyArgPheGlnValSer	1617
Db	2679	CTGTCTTAGCACTACTTCAAGAGCTGGTGGTTTAAAGTGGGACATTTCAAGTTCT	2738
Qy	1618	ValAlaAlaAspGlyAlaGlnLysGlnGlyLysAsnLysSerGlnAspAlaLysSerVal	1637
Db	2739	GTTGCAGACACAGGTGCTCCAGAAAGAGGTAAATAAGTCAGAAAGATGCAAACTCTGT	2798
Qy	1638	HisPheGluSerSerThrSerGluSerSerValLeuSerSerSerSerProGluSerThr	1657
Db	2799	CATTGTGAATCCAGACCTCAGAGTCTCTCAGTCTATCAAGTATGATCCAGAGAGTACC	2858
Qy	1658	LeuValLysProGluProAsnGlyIleThrIleProGlyIleSerSerAspValProGlu	1677
Db	2859	TTGATGAACACAGAGCCGAATGGATATACCATCTCTGGATATCTTCTTAGATGGCCAG	2918
Qy	1678	SerAlaHisLysThrThrAlaSerGlnAlaLysSerAspThrGlnProThrLysVal	1697
Db	2919	AGTGCCCAAAAATACTAGCTCAGAGGCAAAAGTCAACACTGGCAGCAGCTCAAAAGTT	2978
Qy	1698	GlyArgPheGlnValThrThrThrAlaAsnLysValGlyArgPheSerValSerLysThr	1717
Db	2979	GGAAGTTTCAAGGAGCACTACAGCAAAACAAAGTGGGTGTTCTCTGTATCAAAAAC	3038
Qy	1718	GluAspLysIleThrAspThrLysLysGlnGlyProAlaAlaSerProPhePheAsp	1737
Db	3039	GAGACAAAGTACTAGCAAAAGAAAGAAAGAACACAGTGGCATCTCTCTTTATGAT	3098
Qy	1738	LeuGlnGlnAlaValLeuProAlaValIleProLysLysGlnLysProGluLeuSerGlu	1757
Db	3099	TTGGAAACAAGCTGTTCTCTCTGCTGTATACAAAGAAABARAAGCTGAAGTCAAG	3158
Qy	1758	ProSerHisLeuAsnGlyProSerSerAspProGluAlaAlaPheLeuSerArgAspVal	1777

Db 3159 CCTCACATCTAATGAGCGCTCTTCTGACCCGAGGCGCTTTTAAAGTAGGATGTG 3218
QY 1778 AspAspGlySerGlySerProHisSerProHisGlnLeuSerLysSerLeuProSer 1797
Db 3219 GATGATGATCCGGTAGTCCACACTCGCCCATGAGCTCAAGAGCCTTCTTAGC 3278
QY 1798 GlnAsnLeuSerGlnSerLeuSerAsnSerPheAsnSerSerTyMetSerSerAspAsn 1817
Db 3279 CAGAACTTAAGTCAAAAGCCTTAGTAATTCCTTTAACTCTTACATGAGTAGCAGACAT 3338
QY 1818 GluSerAspIleGluAspGluAspLeuLysLeuGlnLeuArgArgLeuArgAspLysHis 1837
Db 3339 GAGTCAGATATCGAAGATGAAGACTTAAGTTAGAGCTGACGACTAGATGAATAAACAT 3398
QY 1838 LeuLysGluIleGlnAspLeuGlnSerArgGlnLysHisGluIleGluSerLeuTyThr 1857
Db 3399 CTCAAAGAGATTGAGGACTGACAGATGCGCAGAGCATGAATGAATCTTTGTATACC 3458
QY 1858 LysLeuGlyLysValProProAlaValIleIleProProAlaAlaProLeuSerGlyArg 1877
Db 3459 AAACCTGGGCAAGGTGCCCCCTGCTGTATTATTCTCCCGAGCTGCTCCCTTTCAAGGAGA 3518
QY 1878 ArgArgArgProThrLysSerLysGlySerLysSerSerArgSerSerSerLeuGlyAsn 1897
Db 3519 AGACGAGCAGCCCACTAAAGCAAGGAGCAAACTTAGTCAGAGAGTTCTTGAGGAGAT 3578
QY 1898 LysSerProGlnLeuSerGlyAsnLeuSerGlyGlnSerAlaAlaSerValLeuHisPro 1917
Db 3579 AAAAGCCCCAGCTTTCAGGTACCTGCTGTCAGATGACGCTTCAGCTTCACACCCC 3638
QY 1918 GlnGlnThrLeuHisProProGlyAsnIleProGlnSerGlyGlnAsnGlnLeuLeuGln 1937
Db 3639 CAGCAGACCTCCACCTCTCTGAGCAACTCCCAAGTCCGAGCAAGATCAGCTGTACAG 3698
QY 1938 ProLeuLysProSerProSerSerSerAspAsnLeuTySerAlaPheThrSerAspGlyAla 1957
Db 3699 CCCCTTAAGCATCTCCCTCCAGTAGACAACCTTATTCAAGCTTCACCAAGTAGATGTGCC 3758
QY 1958 IleSerValProSerLeuSerAlaProGlyGlnGlyThrSerSerThrAsnThrValGly 1977
Db 3759 ATTTCAGTACCAAGCCTTTCGCTCCAGTCAAGGACCAAGACCAAGACCAACACTGTGGG 3818
QY 1978 AlaThrValAsnSerGlnAlaAlaGlnAlaGlnProProAlaMetThrSerSerArgLys 1997
Db 3819 GCACAGTGAACAGCAGCAGCCGCCCAAGCTCAGCTCCTGCCATGACCTCCAGCAGAGAG 3878
QY 1998 GlyThrPheThrAspAspLeuHisLysLeuValAspAsnTyrAlaArgAspAlaMetAsn 2017
Db 3879 GGCACATTCAAGATGACTTGCAACAAGTTGGTAGACAATGGCCCGAGATGCCATGAT 3938
QY 2018 LeuSerGlyArgArgGlySerLysGlyHisMetAsnTyrGlyGlyProGlyMetAlaArg 2037
Db 3939 CTCACAGGAGAGAGAGCAAGGAGCAATGAATTAMGAGGGCCCTCGAATGGCAAGG 3998
QY 2038 LysPheSerAlaProGlyGlnLeuCysIleSerMetThrSerAsnLeuGlyGlySerAla 2057
Db 3999 AAGTCTCTGACCTGGGCAACTGTGACTCTCCATGACTCGAAGCTGGGTGGCTTGCC 4058
QY 2058 ProIleSerAlaAlaSerAlaThrSerLeuGlyHisPheThrLysSerMetCysProPro 2077
Db 4059 CCCATCTCTGAGATCAGTACTCTCTAGTCACTTCAACAAGTATATGTCCCCCA 4118
QY 2078 GlnGlnTyrGlyPheProAlaThrProPheGlyAlaGlnTyrPserGlyThrGlyGlyPro 2097
Db 4119 CAGCAGATAGGCTTTCAGTACCCCAATTTGGCGCTCAATGAGTGGGACGGGTGGCCCA 4178
QY 2098 AlaProGlnProLeuGlyGlnPheGlnProValGlyThrAlaSerLeuGlnAsnPheAsn 2117
Db 4179 GCACCAAGCCACTTGGCAGATTCCAACTGTGGAACTGCTCTTGACAAATTTCAAC 4238
QY 2118 IleSerAsnLeuGlnLysSerIleSerAsnProProGlySerAsnLeuArgThrThr 2136
Db 4239 ATCAGCAATTTGAGAAATCATCAGCAACCCCCAGGCTCCAACTGTGGGAGCACT 4295

Search completed: September 25, 2004, 13:36:14
Job time : 1955 secs